

Leveraging Wearable Sensors for Real-World Monitoring and Medication Response Evaluation of Parkinson's tremor

Diogo C. Soriano^{1,2}, Yordan Raykov³, Nienke A. Timmermans⁴, Timothy West⁵, Erik Post⁴, Max Little⁶, Rick C. Helmich⁴, Bastiaan R. Bloem⁴, Luc J.W. Evers⁴, Hayriye Cagnan^{5*}

¹MRC Brain Network Dynamics Unit, Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, United Kingdom.

²Center for Engineering, Modeling and Applied Social Sciences (CECS), Federal University of ABC, São Bernardo do Campo, Brazil.

³School of Mathematical Sciences, University of Nottingham, Nottingham, United Kingdom.

⁴ Center of Expertise for Parkinson and Movement Disorders, Department of Neurology, Donders Institute for Brain, Cognition and Behavior, Radboud University Medical Center, Nijmegen, The Netherlands.

⁵Department of Bioengineering, Imperial College London, London, United Kingdom.

⁶Department of Computer Science, University of Birmingham, Birmingham, United Kingdom.

*Corresponding author:

Hayriye Cagnan

h.cagnan@imperial.ac.uk

Abstract

Wearable technologies offer powerful means for objectively and accurately assessing symptom severity, disease progression, and medication efficacy in Parkinson’s disease over extended periods. Despite their potential, real-world deployment for long-term monitoring remains challenging due to the absence of a ground truth and variability introduced by everyday activities, which obscure the extraction and interpretation of clinically meaningful digital biomarkers. In this study, we introduce a novel approach for characterizing rest tremor fluctuations in the context of the Personalized Parkinson Project, an initiative aimed to unveil disease characteristics through longitudinal monitoring. The proposed method is validated against clinical scores in a large cohort (N = 477 at baseline; N = 418 at one-year follow-up) and applied to quantify rest tremor variability across three patient groups: those not taking medication, those unresponsive to levodopa, and those responsive to levodopa. We then introduce a novel metric to capture the strength of *repetitive* tremor fluctuations, which shows a significant correlation with the levodopa equivalent daily dose (LEDD) in patients experiencing motor complications. This relationship between rest tremor dynamics, off-state clinical scores, and medication variables—captured during passive monitoring—offers a promising tool for improving clinical decision-making and disease management in Parkinson’s care.

Introduction

Parkinson's disease (PD) is the fastest-growing neurological disorder [Dorsey et al., 2018; Ou et al., 2021], characterized by a range of motor symptoms — such as tremor, bradykinesia, and rigidity — as well as non-motor symptoms including depression, anxiety, sleep disturbances, and cognitive decline [Bloem et al., 2021]. Currently, diagnosis and long-term monitoring rely on clinic-based assessments [Rovini et al., 2017; Tolosa et al., 2021], which can vary significantly across visits, hospitals and clinicians and can only provide a static window into a patient's experience of the disease.

Wearable sensors offer a promising and cost-effective solution for enhancing clinical care by providing objective biomarkers that can support early diagnosis, track daily symptom variations, monitor disease progression, and evaluate medication effectiveness [Maetzler et al., 2013; Rovini et al., 2017; Moreau et al., 2023; Virbel-Fleischman et al., 2023; Varghese et al., 2024]. Harnessing symptom variability is particularly relevant for PD tremor, which is a highly variable symptom that waxes and wanes spontaneously [Dirkx et al., 2023] and is markedly influenced by stress [Dirkx et al., 2020] as well as movements of other limbs [Zach et al., 2015]. As a consequence, brief "snapshots" obtained during laboratory or clinical evaluations of differ from at-home measurements of tremor [Burq et al., 2022]. Furthermore, unlike bradykinesia and rigidity, PD tremor does not respond to even high doses of levodopa in 39% of patients [Dirkx et al., 2019; Zach et al., 2020] and dopamine-resistance may dynamically change during the disease course [Swinnen et al., 2025] and is enhanced under cognitive stress [Zach et al., 2017]. Hence, tracking whether tremor is dopamine-resistant is relevant for optimal disease management.

While substantial progress has been made in developing devices, power-efficient sensors, and data processing techniques to derive digital biomarkers for tremor [Rovini et al., 2017; Sigcha et al., 2021], passive monitoring under free-living conditions remains a significant challenge [Del Din et al., 2016; Sica et al., 2021]. The variability in physical activities, physiological states, and environmental conditions (e.g., stress levels, sleep quality, and differing work and home environments) within and between patients, introduces complexities in interpreting measurements derived from free-living conditions [Jung et al., 2020].

We hypothesized that isolating repetitive patterns in patients' tremor profiles across multiple days enables the extraction of key variables that capture fluctuations in response to medication to support clinical management. Objective and precise measurements of tremor fluctuations could not only enhance clinical decision-making but also empowe individuals with PD to gain deeper insights into their symptoms and how they respond to treatment.

This study presents an efficient and computationally inexpensive approach to evaluate daily fluctuations in rest tremor, leveraging data from the Personalized Parkinson Project (PPP) [Bloem et al., 2019]. The PPP is a large-scale initiative combining clinical assessments with long-term passive monitoring of 520 PD patients across two years, establishing a robust foundation for the longitudinal characterization of PD through wearable technology. Our algorithm, trained and validated on an independently labelled dataset [Evers et al., 2020], was

1 applied to the PPP to investigate recurring variations in rest tremor across days in real-world
2 settings.

3 The utility of symptom monitoring under free-living conditions is validated by demonstrating
4 an alignment of wrist-derived tremor measures — probability and severity — with in-clinic
5 Unified Parkinson’s Disease Rating Scale (UPDRS) scores. We introduce a frequency-based
6 approach to identify *repetitive* patterns, i.e., fluctuations in rest tremor that systematically
7 repeat over days, and explore their relationship with medication schedules across three
8 groups: (1) non-medicated patients; (2) medicated patients classified as levodopa non-
9 responsive; and (3) medicated patients classified as levodopa responsive. Our findings
10 revealed differences between groups, highlighting the feasibility of isolating fluctuations in
11 tremor under free-living conditions. Our proposed measure, assessing the strength of the
12 repeated tremor fluctuations, correlated with the levodopa equivalent daily dose (LEDD) and
13 the time spent in the OFF state (UPDRS 4.3), a composite measure of the amount of time spent
14 with PD symptoms. The ability to quantify rest tremor fluctuations in free-living conditions
15 and their relationship with medication efficacy and dose scheduling represents a valuable tool
16 for improving clinical management and personalized treatment strategies in PD.

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

Results

Tremor detection model and statistical association with clinical scores

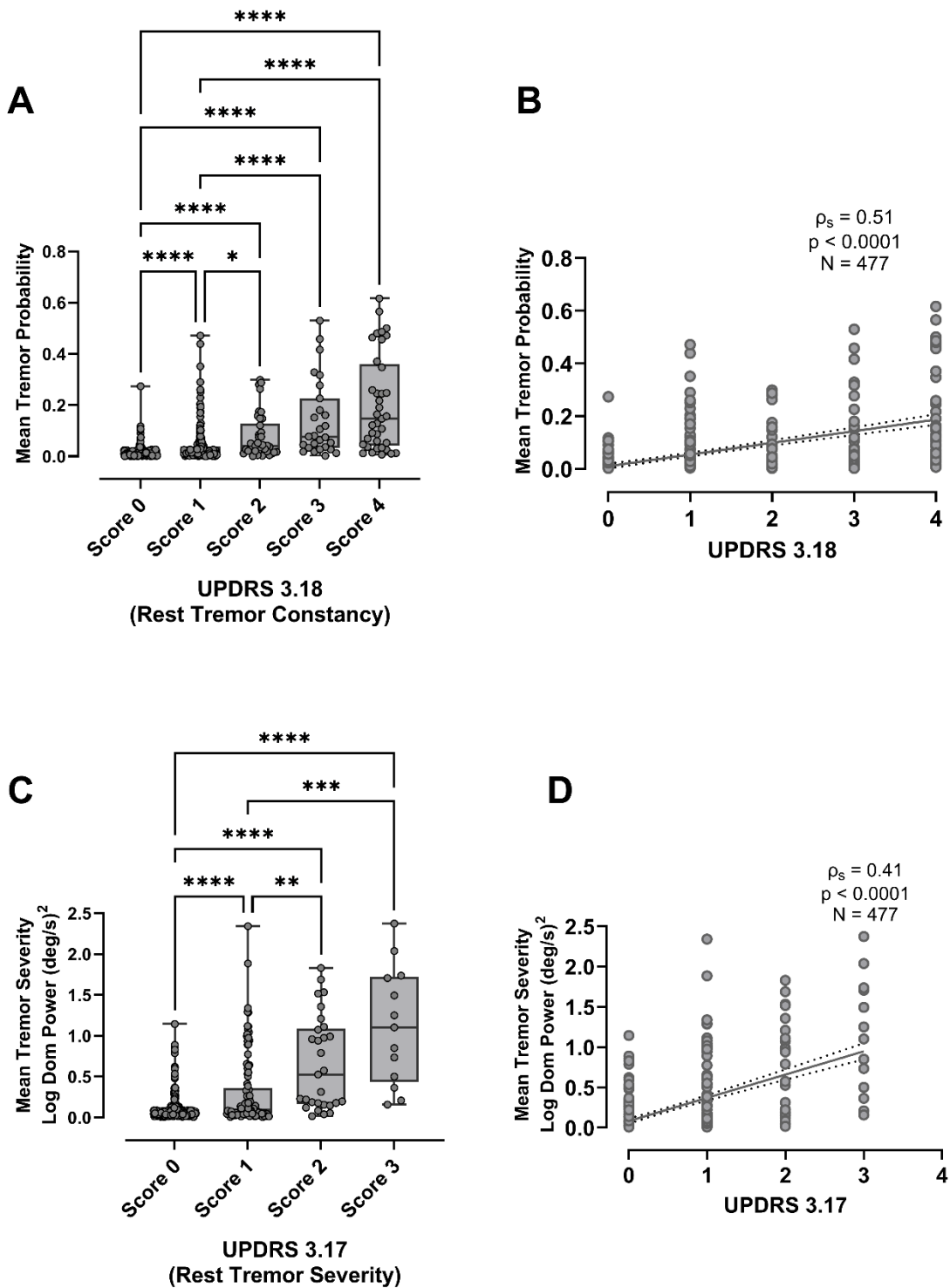
Our rest tremor detection approach combined two classifiers: one for detecting tremor (*tremor detection*), and the other to establish the presence or absence of significant arm activity (*rest condition identification*). Tremor and arm-activity detectors were trained on a manually labelled dataset capturing unscripted daily life activities, the *Parkinson@Home Validation study* [Evers et al., 2020]. Briefly, standardized clinical assessments and wrist sensor recordings during unscripted free-living activities were obtained from a group of 25 PD patients in their home environment. The assessors encouraged the participants to perform indoor and outdoor habitual activities for at least 1 hour, both before and after intake of dopaminergic medication. Video recordings were then labelled by trained research assistants. Annotated data from seven patients with tremor were used to train two logistic regression classifiers, one for tremor and the other for arm activity.

Four types of features were tested for optimal tremor and arm activity detection: 1) power spectral density (PSD) based features [Heida et al., 2013; Luft et al. 2019; Hssayeni et al., 2019; McGurrin et al., 2021], 2) Mel-frequency Cepstral Coefficients (MFCCs) [San-Segundo et al., 2016, San-Segundo 2020, Sigcha et al., 2021], 3) data-range (statistical quantiles) based features [Hammerla et al., 2013] and 4) root mean square (RMS) attributes [Patel et al., 2009] as detailed in the Supplementary Material.

Our analysis (Figure **S2**, Supplementary Material) revealed that combining acceleration and angular velocity yielded no additional predictive power for detecting tremor and arm activity. Moreover, MFCCs provided the best feature set, delivering performance comparable to using all four features combined. Therefore, the final tremor detection model employed MFCC features extracted from gyroscope recordings only. This approach avoided computationally expensive gravitational corrections and resulted in a relatively small feature set. When validated on the PD@Home dataset using leave-one-subject-out cross-validation, the final tremor detector achieved a sensitivity of 0.65 ± 0.18 with a specificity fixed at 0.95 ± 0.00 , yielding an AUC of 0.91 ± 0.06 . The arm activity detector achieved a sensitivity of 0.74 ± 0.17 , a specificity of 0.78 ± 0.15 , and an AUC of 0.88 ± 0.16 (Supplementary Material).

The tremor detection algorithm was then applied to the unlabelled PPP dataset acquired during unscripted free-living activities. Figure **1** shows the agreement between our wrist-derived measures - the mean tremor probability and the mean tremor severity - with the corresponding clinical scores (rest tremor constancy - UPDRS 3.18, and rest tremor severity - UPDRS 3.17) considering eight weeks of data acquired after the first clinical visit. Both mean tremor probability and severity significantly differed when stratified according to the clinical scores (Fig. **1A**: Kruskal-Wallis' test: $H = 133.1$, $p < 0.0001$; and Fig. **1C**: Kruskal-Wallis' test: $H = 92.59$, $p < 0.0001$), as indicated by the post hoc multiple comparison tests corrected by Dunn's method. Wrist-based measures also significantly correlated with the corresponding clinical scores (Figs. **1B** and **1D**, mean tremor probability and UPDRS 3.18: $\rho_s = 0.51$, $p < 0.0001$, $N = 477$; mean tremor severity and UPDRS 3.17: $\rho_s = 0.41$, $p < 0.0001$, $N = 477$). Comparable results were obtained when analysing from the second clinical visit ($N = 418$) considering eight

1 weeks of passive monitoring acquired around the assessment (Figure S6 – Supplemental
 2 Material).



3

4 Figure 1: (A) Boxplots for the wrist-estimated average tremor probability stratified into ascending rest
 5 tremor constancy scores (UPDRS 3.18) and (B) the respective dispersion, linear regression, and
 6 Spearman correlation value; (C) Boxplots for the wrist-estimated mean tremor severity stratified in

1 ascending rest tremor clinical severity scores (UPDRS 3.17) and **(D)** the respective dispersion, linear
2 regression, and Spearman correlation value. Data from the first clinical assessment.

3

4 **Reappearing tremor fluctuations and relation with medication variables**

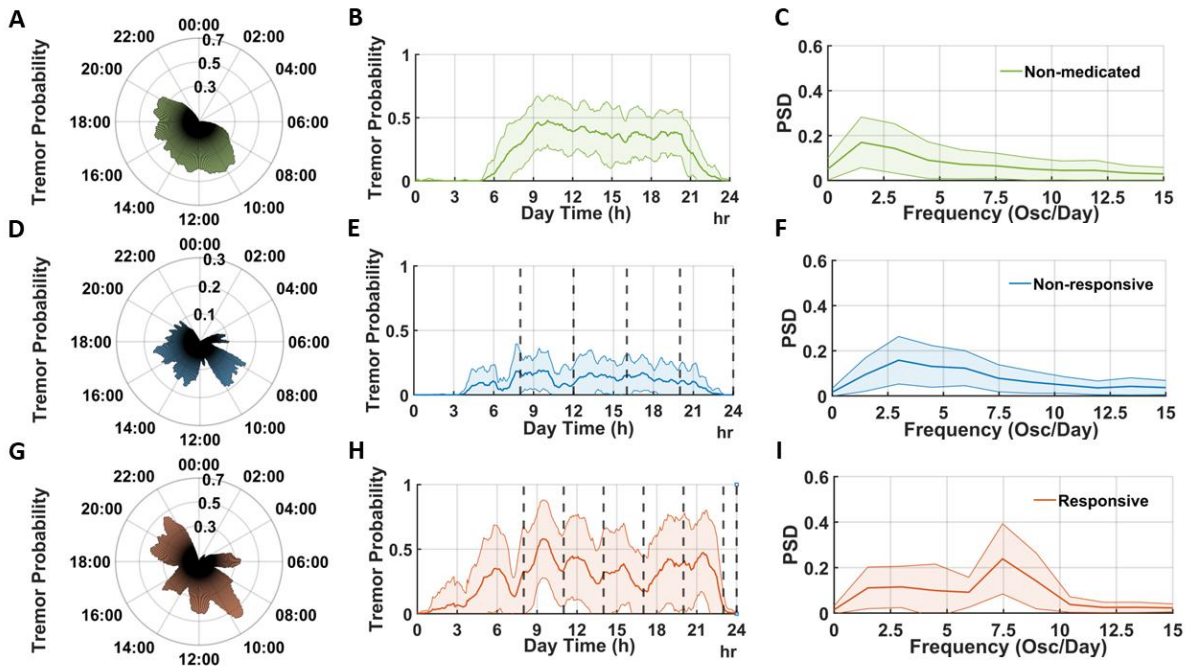
5 We hypothesized that evaluating periodic tremor fluctuations across days could provide
6 means to assess key clinical variables associated with medication effectiveness. To investigate
7 this, we analysed three different cohorts in the context of medication intake: (1) non-
8 medicated group; (2) medicated group taking levodopa at least three times a day and clinically
9 rated as non-responsive to medication according to the levodopa challenge; (3) medicated
10 group taking levodopa at least three times a day and clinically rated as responsive to
11 medication according to the levodopa challenge.

12 Figs. **2A-C** show the polar histogram for the average tremor probability across days, its average
13 \pm standard deviation (shaded – Fig. **2B**), and its average \pm standard deviation power spectral
14 density for a typical non-medicated patient. Sleeping intervals can be inferred from the
15 evening periods with no tremor activity (Figs. **2A** and **2B**), followed by an increase in tremor
16 probability in the morning, the sustained tremor level during the daytime and the evening
17 decline. This pattern contributes to a dominant component at the first frequency coefficient
18 (approximately 1-1.5 oscillations/ day) followed by a typical exponential monotonic decay at
19 higher frequencies (Fig. **2C**). Analogous panels are presented for a typical levodopa non-
20 responsive patient (Figs. **2D**, **2E**, and **2F**) and a levodopa responsive patient (Figs. **2G**, **2H**, and
21 **2I**). In general, levodopa non-responsive patients tend to exhibit broader tremor fluctuations
22 across days compared to the non-medicated group, reflected in additional frequency
23 components in the PSD, which broaden and shift the primary peak observed around 1–1.5
24 oscillations per day in non-medicated patients. In contrast, levodopa responsive patients
25 typically demonstrate well-defined oscillatory modes (Fig. **2H**) with shifted and enhanced
26 frequency components (Fig. **2I**).

27

28

29

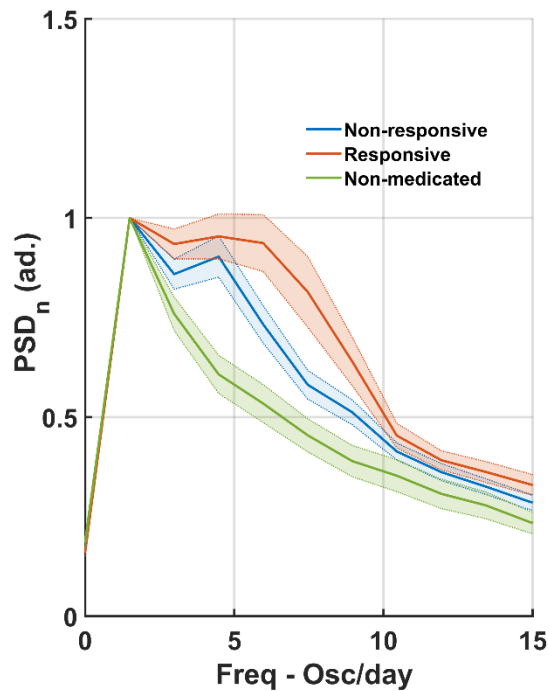
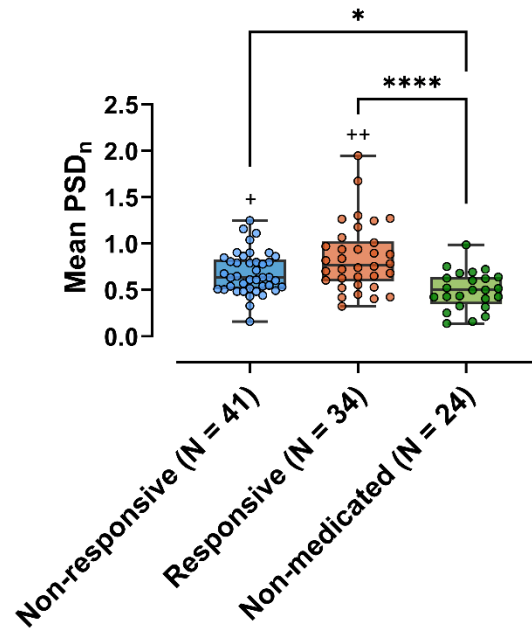


1

2 Figure 2: (A) Wrist-estimated average cycle of the daily circadian tremor probability (N = 46 days) for a
 3 typical non-medicated patient; (B) Mean and standard deviation (shaded) of tremor probability
 4 fluctuations across days; (C) Mean power spectral density (PSD) function for daily tremor fluctuations
 5 and its standard deviation (shaded) normalized by the sum of the coefficients; (D, E, F) – analogous
 6 panels for a typical levodopa non-responsive patient (N = 35 days). Vertical traced lines denote the
 7 prescribed medication intake starting times; (G, H, I) – analogous panels for a typical patient responsive
 8 to levodopa medication (N = 26 days).

9

10 Figure 3A illustrates the average power spectral density normalized (within patient) by the
 11 first oscillatory component, defining the PSD_n for wrist-estimated tremor probability
 12 (Methods Section). The non-medicated group exhibited an exponential decay with no
 13 coherent tremor activity, whereas the medicated cohorts exhibited regular rhythmic
 14 components compatible with the typical daily dose schedule (3-10 oscillations/day)
 15 superimposed over the exponential background. This effect was further exaggerated in the
 16 levodopa responsive group. A significant difference was observed in the mean PSD_n
 17 coefficients within the medication intake range across the groups (Kruskal-Wallis's test: $H =$
 18 18.47 , $p < 0.0001$, Fig. 3B). Post-hoc Dunn's tests revealed that both levodopa responsive and
 19 non-responsive cohorts exhibited higher PSD_n averages compared to the non-medicated
 20 group.

A**B**

1

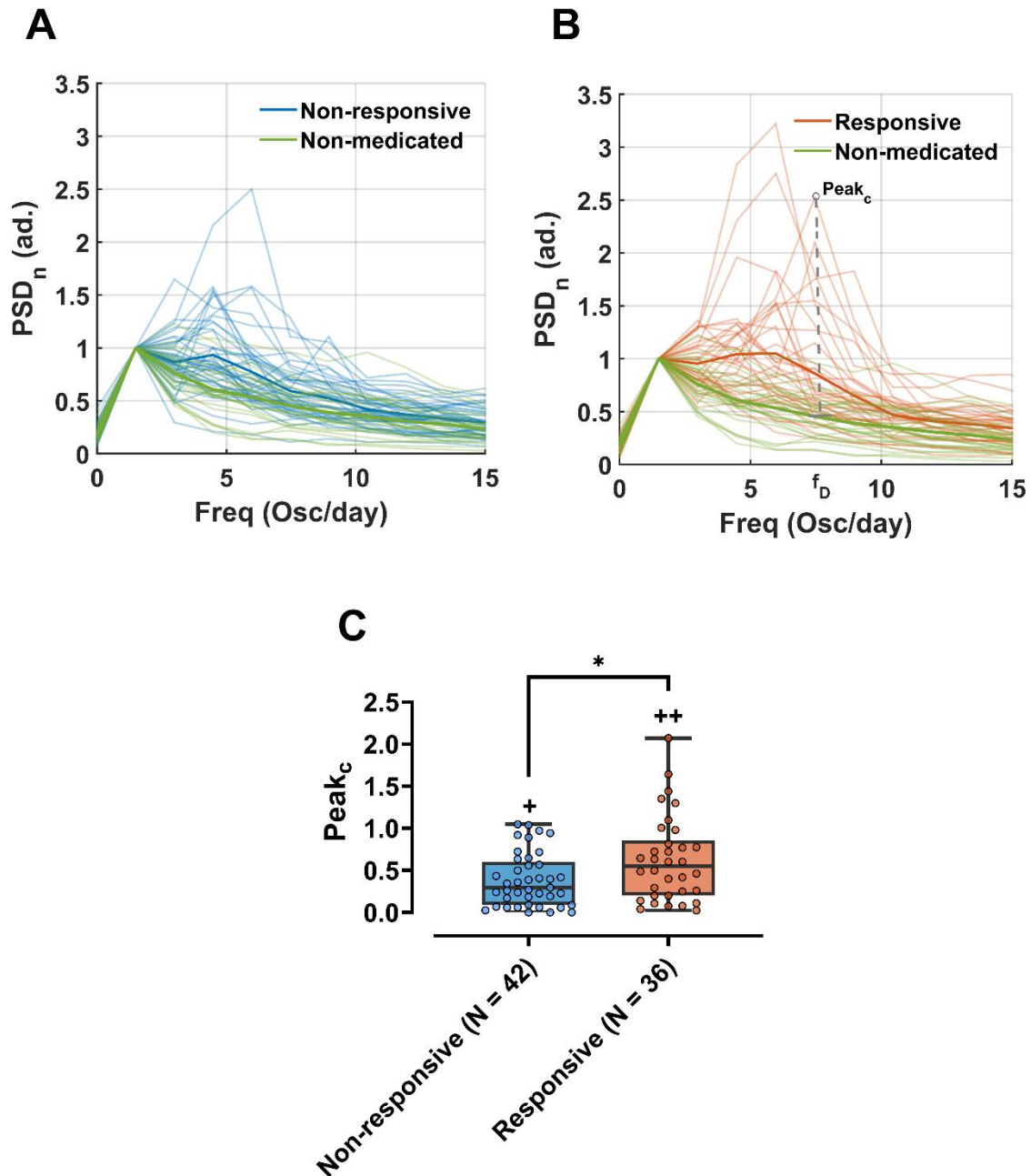
2 Figure 3: (A) Average power spectral density of daily sensor-estimated tremor probability normalized
 3 by the first oscillatory component (PSD_n). Continuous traces and shaded regions show, respectively,
 4 the PSD_n average and standard error of the mean for each cohort: levodopa non-responsive, levodopa
 5 responsive, and non-medicated. Stratification by medication-responsiveness is performed according
 6 to tremor constancy clinical score (UPDRS 3.18) changes observed under the levodopa challenge. (B)
 7 Boxplot for the mean PSD_n within the medication (schedule) intake frequency range ([3-10]
 8 oscillations/day) for the cohorts. Patients from the first and second clinical assessments were
 9 considered as independent samples.

10

11 Figure 4 (A and B) present the individual normalized PSDs for wrist-estimated tremor
 12 probability across the investigated cohorts, along with their respective averages (thick line).
 13 Both the levodopa non-responsive and responsive groups contained patients with enhanced
 14 frequency components clearly different from the monotonic decay observed for the non-
 15 medicated group as shown in both panels. These enhanced frequency components occur
 16 between 3 and 10 oscillations per day. Notably, the frequency peaks are more prominent in
 17 the levodopa-responsive cohort (Fig. 4B) but also present in the non-responsive group.
 18 Indeed, the strength of the consistent tremor fluctuation ($Peak_c$) – defined as the difference
 19 between patient's PSD peak and the respective PSD average coefficient for the non-medicated
 20 population (Fig. 4B) – is significantly higher for the levodopa responsive group when
 21 compared to the non-responsive group (Mann-Whiney test: $U = 495$, $p = 0.03$) (Fig. 4C).

22

23



1

2 Figure 4: (A) Individual PSD_n (shaded) and group average (thick line) for the levodopa non-responsive
 3 (blue) and non-medicated (green) cohorts; (B) Individual PSD_n (shaded) and group average (thick line)
 4 for the levodopa medication-responsive (red) and non-medicated (green) cohorts. Peak_c illustrates
 5 the estimation of the strength of the consistent tremor fluctuations across days based on patient PSD_n
 6 peak corrected by the non-medicated average. f_D denotes patient's dominant frequency in the
 7 medication intake frequency range; (C) Peak_c boxplot for the levodopa non-responsive and levodopa
 8 responsive cohorts.

9

10

1 Given the heterogeneous behavior of Peak_c observed within the clinically defined groups, it is
 2 reasonable to suggest that different clinical factors may also influence the strength of tremor
 3 fluctuations, which supports an exploratory analysis. Table 1 shows the pairwise correlations
 4 between Peak_c and key clinical variables involving medication information, disease
 5 progression and arm activity, summarized by (Methods Section): **Resp** - clinical responsiveness
 6 to medication; **Wearing OFF Quest** improvement of... in daily life after medication intake
 7 according to the wearing OFF questionnaire; **UPDRS 4.3** - time spent in OFF state; **Disease**
 8 **Duration** – disease duration in months; **LEDD** - Levodopa equivalent daily dose; **Doses**: the
 9 number of levodopa daily doses; **Age**: patients' age; **ArmActv**: mean detected arm activity;
 10 **PaidJob** - binary variable for paid activity. Interestingly, the correlations obtained suggest that:
 11 (1) arm-activity, which is expected to be affected by the daily routine tasks, doesn't relate to
 12 Peak_c or any other variable; (2) The time spent in OFF state correlates both with the Peak_c,
 13 LEDD and number of levodopa doses; (3) LEDD correlates with Peak_c, clinically rated
 14 medication responsiveness, number of levodopa doses, time spent in OFF state and disease
 15 duration.

16 **Table 1:** Spearman correlation between patients' clinical variables and the Peak_c evaluated
 17 from the individual tremor probability. Significant positive and negative correlations are
 18 highlighted in red and blue, respectively. **Resp**: clinical tremor responsiveness to levodopa
 19 medication; **Wearing OFF Quest**: tremor improvement with medication - Wearing OFF
 20 questionnaire (yes: 1; no: 0); **UPDRS 4.3**: time spent in levodopa OFF state; **Disease Duration**:
 21 disease duration in months; **LEDD**: Levodopa equivalent daily doses (mg); **Age**: patient's age
 22 in years; **ArmActv**: mean arm activity probability; **PaidJob** – patient paid job activity (yes:1;
 23 no: 0). P-values were adjusted for a false discovery rate of 0.05 according to Benjamini &
 24 Hochberg, 1995 procedure.

25

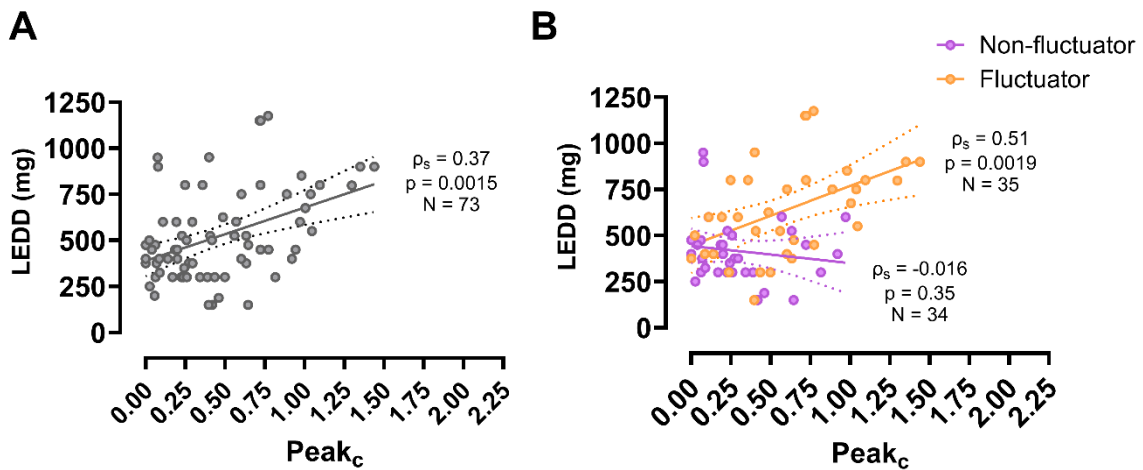
	Resp	Wearing OFF Quest	UPDRS 4.3	Disease Duration	LEDD	Doses	Age	ArmActv	PaidJob	Peak _c
Resp	-	0.34	0.35	0.19	0.48	0.29	-0.19	-0.12	0.03	0.29
Wearing OFF Quest	0.34	-	0.30	0.08	0.28	0.23	-0.23	-0.02	0.14	0.26
UPDRS 4.3	0.35	0.30	-	0.20	0.51	0.46	-0.17	0.09	0.02	0.43
Disease Duration	0.19	0.08	0.20	-	0.50	0.26	0.14	-0.06	-0.21	0.26
LEDD	0.48	0.28	0.51	0.50	-	0.53	-0.07	0.10	-0.24	0.40
Doses	0.29	0.23	0.46	0.26	0.53	-	-0.13	0.16	-0.14	0.34
Age	-0.19	-0.23	-0.17	0.14	-0.07	-0.13	-	-0.15	-0.63	-0.07
ArmActv	-0.12	-0.02	0.09	-0.06	0.10	0.16	-0.15	-	0.07	0.05
PaidJob	0.03	0.14	0.02	-0.21	-0.24	-0.14	-0.63	0.07	-	-0.15
Peak _c	0.29	0.26	0.43	0.26	0.40	0.34	-0.07	0.05	-0.15	-

26

27

28

1 As both LEDD and UPDRS 4.3 correlate with $Peak_c$, it is reasonable to postulate that the
 2 experience of OFF states (UPDRS 4.3 ≥ 1) contributes to the consistent tremor fluctuations
 3 observed. Figure 5 further illustrates that the correlation between $Peak_c$ and LEDD (Fig. 5A) is
 4 primarily driven by the subset of patients experiencing OFF states (symptom fluctuator cohort,
 5 UPDRS 4.3 ≥ 1). In contrast, patients without motor fluctuations (UPDRS 4.3 = 0) show no
 6 correlation between $Peak_c$ and LEDD (Fig. 5B). This stratification based on UPDRS 4.3 offers a
 7 more robust explanation ($\rho_s = 0.51$, $p = 0.002$) for the population-level relation between $Peak_c$
 8 and LEDD observed in Fig. 5A ($\rho_s = 0.37$, $p = 0.001$).



9

10 **Figure 5:** (A) Levodopa equivalent daily dose (LEDD) vs. the strength of the consistent tremor
 11 fluctuations across days ($Peak_c$) for the medicated cohort evaluated from tremor probability in the
 12 levodopa constancy responsiveness scenario. Outliers were excluded for both variables. Spearman
 13 correlations (ρ_s), p-values (p) and number of samples (N) are shown in the panels. (B) LEDD vs. $Peak_c$
 14 dispersion considering further stratification into symptom fluctuator (UPDRS 4.3 ≥ 1) and non-
 15 fluctuator (UPDRS 4.3 = 0) groups. Fewer patients ($N = 69$) had UPDRS 4.3 data available, explaining
 16 the sample size difference compared to panel (A).

17

18 Finally, the sub-group of patients experiencing motor-fluctuations showed a marginal
 19 correlation between the dominant frequency (f_D – Fig. 4B) and the number of levodopa doses
 20 taken ($\rho_s = 0.28$, $p = 0.09$, $N = 38$). A key limitation in associating these variables lies in the
 21 minimal signal to noise ratio required to accurately describe this relationship, as patients with
 22 low $Peak_c$ may contribute to spurious f_D values. Alternatively, if we select a group with higher
 23 $Peak_c$ regardless of the initial clinical levodopa responsiveness stratification, which can be
 24 performed using a clustering approach (k-means), a significant correlation emerged between
 25 f_D and the number of levodopa doses ($\rho_s = 0.46$, $p = 0.03$, $N = 22$), which once again highlights
 26 the relationship between the tremor fluctuation strength and the medication variables.

27 Overall, the strength of consistent fluctuations across days based on wrist-estimated severity
 28 (i.e., the tremor amplitude) provided comparable inferences to those based upon tremor
 29 probability reported in Figures 2-5 and Table 1. Additional details are provided in the
 30 Supplementary Material (see Section Tremor fluctuation dynamics – Tremor Severity). In the
 31 group analysis of PSD_n under the severity scenario (Fig. S7 – Supplementary Material), the

1 differential effect between levodopa-responsive and non-responsive groups was smaller than
2 that observed for tremor probability (Fig. 3). Furthermore, unsupervised clustering of Peak_c
3 derived from wrist-based severity did not reveal a significant relationship between f_D and the
4 number of doses. Confounding factors that influence tremor severity, such as stress or arousal,
5 may introduce additional sources of unexplained variability that may explain the lower effect
6 sizes observed for inferences derived from tremor severity.

7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33

1

2

Discussion

3 This study presents a novel framework for analyzing rest tremor fluctuations in Parkinson's
4 disease (PD) patients under free-living conditions. Methodologically, we introduce an efficient
5 and computationally inexpensive model for detecting rest tremors based on a wrist-worn
6 sensor, and propose a frequency-based metric to capture consistent tremor fluctuations
7 across multiple days. Clinically, wrist-based measures of tremor probability and severity
8 demonstrate significant correlations ($\rho_s \approx 0.5$) with corresponding UPDRS scores assessed
9 during both clinical visits. Additionally, our frequency-based measure for the strength of daily
10 fluctuations in tremor ($Peak_c$) identified consistent patterns that effectively distinguished
11 levodopa-responsive from non-responsive groups. This measure significantly correlated with
12 the levodopa equivalent daily dose (LEDD) in patients experiencing OFF states (UPDRS 4.3 \geq
13 1) suggesting that it is a proxy of medication-induced tremor fluctuations. Furthermore,
14 unsupervised clustering of medicated patients based on $Peak_c$ revealed a group with more
15 consistent tremor fluctuations, whose dominant frequencies were significantly correlated
16 with the number of levodopa doses.

17 Although there is an extensive body of work on detecting PD symptoms using wearable
18 sensors [Maetzler et al., 2013; Rovini et al., 2017; Moreau et al., 2023], relatively few studies
19 focus on continuous monitoring in real-world settings, which poses significant validation
20 challenges [Sigcha et al., 2021]. Historically, algorithm validation has relied on labels derived
21 from episodic and remotely administered motor tests [Lipsmeier et al., 2022; Burq et al., 2022;
22 Oyama et al., 2023], self-reported motor states [San-Segundo et al., 2020], or comparisons
23 with scores obtained during clinical visits [Mahadevan et al., 2020; Adams et al., 2023]. In this
24 study, our metrics demonstrated significant agreement with corresponding clinical scores.
25 Notably, cohorts with moderate to severe rest tremor constancy or severity showed sensor-
26 derived metrics that were significantly distinct from those with mild or no tremor, consistent
27 with findings in [Mahadevan et al., 2020].

28 Previous studies [Burq et al., 2022; Moreau et al., 2023; Oyama et al., 2023] have
29 demonstrated that digital measurements from smartwatches are sensitive to changes
30 associated with medication states, showing small to medium effect sizes during virtually
31 conducted motor assessments in both hospitalized and home-based settings under levodopa
32 OFF and ON conditions. Similarly, [Heijmans et al., 2019] utilized episodic interventions to
33 assess the states of PD patients in daily life, presenting a case study in which inertial signals
34 were leveraged to detect the levodopa OFF state (AUC = 0.73), supporting the feasibility of
35 OFF state detection in such scenarios.

36 Continuous monitoring of levodopa-induced tremor decay has been explored by [Pulliam et
37 al., 2018] using the Kinesia motor assessment system (Great Lakes NeuroTechnologies,
38 Cleveland, OH) during daily activities in home-simulated environments, further advocating for
39 the practicality of tracking tremor fluctuations in real-world settings. Additionally, the
40 Parkinson's KinetiGraph System (PKG, Global Kinetics Corporation, Melbourne, Australia) has
41 proven effective in distinguishing levodopa ON and OFF states based on sensor-derived

1 bradykinesia scores, which correlated moderately to strongly with patient diaries [Ossig et al.,
2 2016]. This system has also been successful in characterizing wearing OFF fluctuations
3 [Farzanehfar et al., 2022]. Finally, [Mahadevan et al., 2020] demonstrated a strong agreement
4 between sensor-derived tremor metrics in levodopa ON and OFF states and the corresponding
5 clinical scores obtained during in-clinic visits.

6 In contrast to previous studies, here, we introduced for the first time a frequency-based
7 approach to assess consistent tremor patterns across multiple days in patients taking
8 levodopa. This approach stands for: 1) Time-shift invariance – ensuring robustness to temporal
9 jitter (e.g., due to variable medication intake time) in the data (Fig. 2); 2) Higher signal to noise
10 ratio – by using ensemble averaging across trials in the frequency domain highlighting
11 components that consistently reappear across days; 3) Normalization by the diurnal cycle –
12 enabling the detection of consistent oscillations relative to other diurnal patterns; 4) Use of a
13 contrasting metric – providing a reference or a control population (non-medicated population)
14 for highlighting the medication effect. This new framework which relied solely on passive
15 monitoring under real-world conditions unveiled stronger frequency components within the
16 levodopa intake range for the medicated group, leading to the development of a measure that
17 correlated significantly with the patients' daily levodopa dosage (LEDD) and the time spent in
18 the OFF state (UPDRS 4.3).

19 By stratifying the medicated patients into two groups according to the UPDRS 4.3 clinical score
20 - those with fluctuating symptoms ($UPDRS\ 4.3 \geq 1$), and those without ($UPDRS\ 4.3 = 0$) -
21 revealed that the relationship between LEDD and tremor fluctuation strength was most
22 pronounced in patients experiencing OFF states ($UPDRS\ 4.3 \geq 1$). This finding suggests that
23 such patients are more likely to exhibit medication-induced tremor reduction, with daily
24 tremor cycles closely tied to their medication schedules. Notably, in the group with greater
25 tremor fluctuation strength, a significant correlation was observed between the dominant
26 frequency of these fluctuations and the number of daily levodopa doses. This confirms the
27 link between average tremor frequency patterns and medication schedules.

28 Although the MDS-UPDRS 4.3 accounts for a broader range of symptomatic worsening to
29 characterize the duration of OFF states, patients experiencing response fluctuations often
30 report increased tremor, [Perepezko et al., 2020]. Furthermore, our analytical framework and
31 inclusion criteria focused on patients exhibiting detectable tremors, prioritizing the objective
32 quantification of tremor changes in fluctuators ($UPDRS\ 4.3 \geq 1$).

33 The observed relationship between LEDD and tremor fluctuation strength in free-living
34 conditions is particularly noteworthy. It suggests that wrist-based monitoring could provide
35 an objective measure of a patient's response to medication, offering valuable insights for
36 tracking disease progression and enhancing clinical management. This is clinically relevant,
37 given that PD tremor does not adequately respond to levodopa in 39% of cases [Zach et al.,
38 2020], while stereotactic surgery (e.g. deep brain stimulation) is highly effective even for
39 dopamine-resistant tremor [Armstrong & Okun, 2020]. Stress can abolish the effect of
40 levodopa on PD [Zach et al., 2017], making it difficult to reliably establish whether PD tremor
41 is dopamine resistant. Hence, earlier detection of dopamine-resistant PD tremor using

1 wearable devices may help clinicians to consider stereotactic treatments in earlier stages of
2 the disease.

3 **Study Limitations**

4 As a main limitation of this work, the proposed approach defines a heuristic with different
5 parameters requiring context-dependent designing options, which may impact its general use
6 and widespread applicability. In addition, the analysis performed here was constrained to a
7 specific population (e.g., patients taking levodopa medication at least three times a day, with
8 no levodopa agonist and exhibiting a minimal level of detectable rest tremor) and a broader
9 investigation would be required for evaluating its potential under different experimental
10 designs.

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38

Methods

Rest tremor detection: a supervised approach

PD participants in the *Parkinson@Home Validation study* [Evers et al., 2020] (n=25) wore lightweight sensors (Physilog® 4, Gait Up S.A., Lausanne, Switzerland) on both wrists. Video recordings were analysed and labelled by trained research assistants, providing annotations for the performed activities (walking, sitting, standing, etc.,) and visible PD symptoms (e.g., tremor).

Data from a triaxial accelerometer (m/s^2) and a gyroscope (deg/s) were acquired at a sampling rate of 200 Hz and downsampled (by decimation) to 100 Hz. Gravity effects on acceleration were filtered out using an L1 trend filter [Kim et al., 2009]. The data was then sliced into 4 s windows with 2 s overlap. Four types of features were tested for optimal tremor and arm activity detection, as detailed in the Supplementary Material. Extracted triaxial features were combined (summed) to provide a single reference-free feature set [Mcgurrin et al., 2021]. Tremor and arm activity detection performances were characterized in terms of the area under the curve (AUC) of the receiver operating characteristics (ROC) for individual feature sets for both acceleration and angular velocity, and for their combination (i.e., all features from all inertial measurements). The final feature set, and the detection performances are summarized in the Results Section with a detailed assessment in the Supplementary Material. Given the superior performance of MFCC features in this labelled data set, this set was taken forward for tremor and arm activity prediction.

The Personalized Parkinson Project

The classifiers trained on the labelled data were taken forward to analyse circadian tremor dynamics and consistent tremor fluctuations in the context of the larger, but unlabelled, Personalized Parkinson Project (PPP) data [Bloem et al., 2019]. This cohort comprises 520 patients followed for 2 up to 3 years with annual in-clinic assessments and continuous daily monitoring under free-living conditions using a wrist-worn sensor device. The in-clinic assessments included an extensive clinical evaluation (e.g., MDS-UPDRS, demographics and lifestyle), the collection of biospecimens, and neuroimaging. In addition, patient completed annual questionnaires e.g. about their motor and non-motor symptoms, sleep disorders, quality of life, and response fluctuations). Continuous monitoring was achieved using a wrist-worn multi-sensor device, the Verily Study Watch (Verily, San Francisco, USA), continuously capturing acceleration, angular velocity, pulse rate, and electrodermal activity. Patients were

1 recommended to use the study watch 24/7 (i.e., all day except for daily charging periods of ~1
2 hour). Further details can be found in [Bloem et al., 2019].

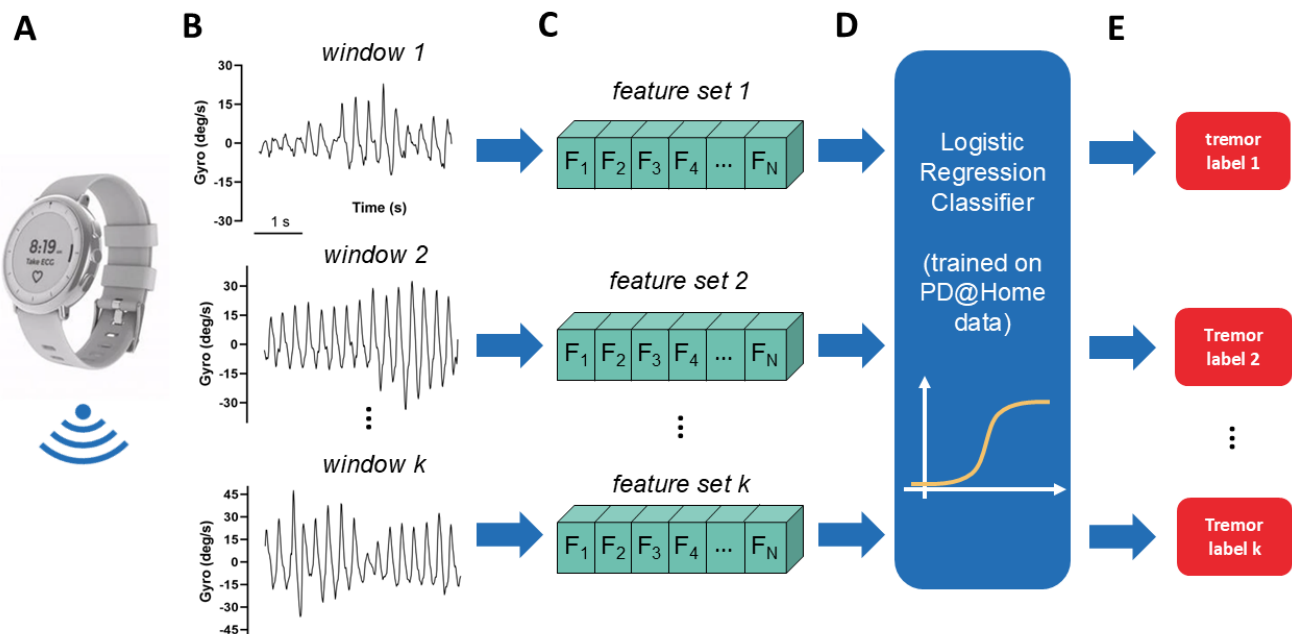
3 **Passive rest tremor monitoring: the Personalized Parkinson Project**

4 The sampling rate was originally set to 50 or 100 Hz. Gyroscope signals were cubic interpolated
5 to ... Hz to ensure that the sampling frequency matched the value set, given hardware
6 sampling jitter. Time series were sliced into 4 s windows, followed by MFCCs feature extraction
7 - the best feature set identified in the supervised scenario. MFCCs were z-score normalized
8 and used as input in the logistic regression classifiers trained on the PD@Home dataset to
9 obtain tremor and arm activity labels in the PPP dataset (Supplementary Material), as
10 schematically represented in Figure 6.

11 Rest tremor probability was quantified as the total number of rest tremor windows detected
12 divided by the total number of rest windows within a 5 min interval. Windows with significant
13 arm activity were not considered in this calculation as also performed in [Mahadevan et al.,
14 2020]. Tremor severity was obtained using $\log_{10}(T_{DP} + 1)$, where T_{DP} is the power in the [3.5
15 – 8 Hz] range considering a bandwidth of 1 Hz around the tremor peak. We opted to use the
16 log scale due to the linear relation between the clinical rating for tremor severity and the log
17 of the tremor oscillation amplitude [Elble et al., 2006]. The argument of the log function was
18 defined to map the tremor severity as zero for windows classified as non-tremor, since T_{DP}
19 would be zero in this condition.

20

21



22

23 **Figure 6:** Schematic representation of the tremor detection framework. (A) The inertial data from the
24 Verily Study Watch is encrypted and sent to the Verily Cloud. (B) stored angular velocity signals were
25 offline interpolated and sliced into windows of four seconds each. (C) The data features (MFCCs) were

1 extracted and used as (D) input to a previously trained logistic regression classifier to obtain (E) a binary
2 tremor label. The same procedure was applied for arm activity detection.

3

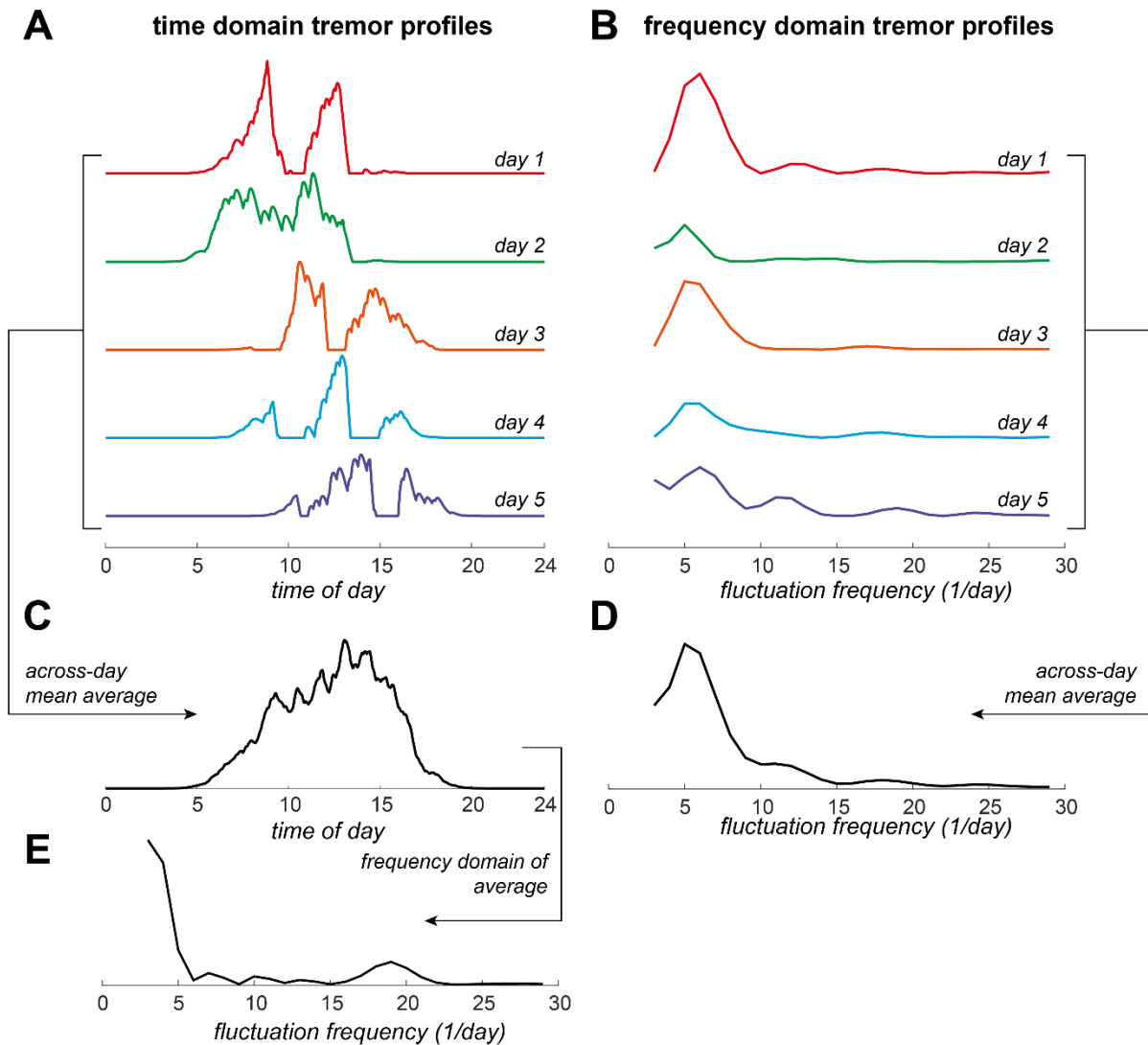
4 **Wrist-based measures and their agreement with clinical scores**

5 Daily sensor data from the eight weeks surrounding the first and second clinical visits
6 (approximately one year apart) were processed using the pipeline outlined in Figure 6 for 477
7 and 418 patients, respectively. For days with at least 60% of data available, the mean rest
8 tremor probability and severity were calculated. These wrist-based measurements were
9 compared to the corresponding clinical scores: rest tremor constancy (MDS-UPDRS 3.18) and
10 severity (MDS-UPDRS 3.17). Patients were stratified into groups based on their respective
11 clinical scores, enabling comparisons between the clinical and wrist-based measures.
12 Spearman correlation was employed to assess the statistical association between the two.

13 **Reappearing rest tremor fluctuations in medicated and non-medicated cohorts**

14 Based on the hypothesis that consistent tremor fluctuations across days may reflect regular
15 deterministic behavior (e.g., daily routines) or external stimuli (e.g., medication schedules),
16 the average daily PSD coefficients were estimated from the wrist-based measures. This
17 approach aimed to capture persistent daily patterns, independent of temporal shifts
18 introduced by variations in the timing of medication intake. Figure 7 demonstrates the
19 approach by analyzing daily tremor probability profiles (Fig. 7A) and their corresponding
20 power spectral densities (PSDs) (Fig. 7B). As the daily PSD coefficients are invariant to time-
21 shifts, averaging across daily observations enhances common frequency components (Fig.
22 7D). In contrast, averaging across the time domain (Fig. 7C) results in the loss of representative
23 daily frequency components due to temporal shifts, as reflected in the resulting PSD shown in
24 Fig. 7E.

25



1

2

3 **Figure 7: (A)** Individual daily circadian tremor probability profiles (simulated for illustration),
 4 and **(B)** their respective power spectral densities. **(C)** Cross-day average in the time domain.
 5 **(D)** Cross-day average in the frequency domain. **(E)** Power spectral density of the across-day
 6 temporal tremor probability average.

7

8 Three different cohorts were analysed: (1) non-medicated patients; (2) patients taking
 9 levodopa at least three times a day, clinically rated as non-responsive to medication; and (3)
 10 patients taking levodopa at least three times a day, clinically rated as responsive to
 11 medication. Patients taking dopamine agonists in addition to levodopa were excluded from
 12 the analysis due to possible complex interactions with levodopa.

13 Levodopa responsiveness was independently assessed by evaluating changes in the severity
 14 and constancy of UPDRS scores between the OFF and ON states. The wrist-based measures
 15 (rest tremor severity and probability) obtained from passive monitoring were used
 16 independently to calculate the average PSD coefficients for the previously described cohorts.

1 While we present levodopa responsiveness based on tremor constancy (probability) in the
 2 main text, analogous results derived from tremor severity are provided in the Supplementary
 3 Material. These two scenarios yield similar inferences, as both rely on the occurrence of rest
 4 tremor, but with severity measures yielding a lower SNR due to increased susceptibility to
 5 confounds of environmental factors and other tremor modulating factors (e.g., stress). Any
 6 similarities or differences between them are explicitly noted in the paper. Data from each
 7 clinical assessment were treated as independent samples. The demographics of the cohorts
 8 used for levodopa responsiveness analysis are summarized in Table 2.

9
 10 **Table 2:** Population demographics considering levodopa responsiveness based on tremor
 11 constancy for levodopa non-responsive (**NR**), and levodopa responsive (**R**), and the non-
 12 medicated (**NM**) cohorts. Averages \pm standard deviation are shown for each variable (except
 13 for gender). LEDD denotes the levodopa equivalent daily dose; H&Y – the Hoehn & Yahr scale
 14 in levodopa OFF condition for assessing disease stage; UPDRS III – the total sum for the Unified
 15 Parkinson’s Disease Rating Scale - part III: motor examination in levodopa OFF condition.

16

COHORT	GENDER (F/M)	AGE (YEARS)	DISEASE DURATION (MONTHS)	LEDD (MG/DAY)	H&Y	UPDRS III
NM	9/15	57 \pm 12.20	31.56 \pm 16.06	N.A.	1.92 \pm 0.41	36.68 \pm 10.03
NR	11/31	63.48 \pm 7.70	36.30 \pm 17.54	457.82 \pm 227.68	2.12 \pm 0.40	38.86 \pm 12.02
R	9/27	62.17 \pm 9.13	46.40 \pm 20.53	651.78 \pm 267.25	2.03 \pm 0.51	43.17 \pm 11.19

17
 18 Daily PSD coefficients were independently estimated from circadian wrist-based variables
 19 (probability and severity) obtained in consecutive 5-minute windows without overlap during
 20 the daily activity interval (06:00–23:00 h). These coefficients were smoothed using a 6-point
 21 (30-minute) moving average filter. Only days with at least 80% of non-missing data within this
 22 interval were included in the daily PSD estimation (details provided in the Supplementary
 23 Material).

24 The PSD coefficients averaged across days were normalized by the first oscillatory component
 25 (lowest non-zero frequency coefficient or fundamental frequency), which represents the
 26 dominant rhythm given by diurnal tremor cycle. This normalization facilitates the analysis of
 27 frequency components relative to the typical strength of rest tremor observed while patients
 28 are awake.

29 Tremor fluctuation strength across days ($Peak_c$) was quantified using a contrast metric to
 30 compare medicated and non-medicated patients. This metric was defined as the difference
 31 between the peak of the average daily normalized PSD coefficients for each patient and the
 32 corresponding non-medicated average normalized PSD coefficients.

33 Since the proposed measure for assessing tremor fluctuations across days may be influenced
 34 by additional factors - such as the amount of arm activity, medication variables (e.g., number
 35 of doses or levodopa equivalent daily dose), and key clinical aspects (e.g., age, disease

1 duration) - we conducted an exploratory analysis by evaluating the correlation matrix,
2 considering: **Resp**: Patient clinical tremor responsiveness to levodopa medication, as defined
3 in [Zach et al., 2020] (Equation 1); **Wearing OFF Quest**: Tremor improvement with medication,
4 assessed via the Wearing OFF questionnaire (yes: 1; no: 0); **UPDRS 4.3**: time spent in levodopa
5 OFF state; **Disease Duration**: disease duration in months; **LEDD**: Levodopa equivalent daily
6 doses (mg); **Age**: patient's age in years; **ArmActv**: mean arm activity probability; **Paid Job** –
7 paid job activity (yes:1; no: 0).

8 **Statistical Analysis**

9 The D'Agostino and Pearson test was used to assess the normality of random variables'
10 distribution. Hypothesis testing was conducted using Student's t-test (t statistic) or ANOVA (F
11 statistic) with Tukey post-hoc corrections for normally distributed data depending on the
12 number of groups. Alternatively, Mann-Whitney (U statistic) or the Kruskal-Wallis' test (H
13 statistic) with Dunn's post-hoc corrections for non-normal distributions. Correlation analyses
14 employed Pearson's correlation coefficient (ρ_p) for normal distributions and Spearman's
15 correlation coefficient (ρ_s) for non-normal distributions.

16 In the boxplots presented, * indicates p-value < 0.05, ** indicates p-value < 0.01, and
17 additional * denotes a 10-fold decrease in the significance level (e.g., p-value < 0.001). When
18 applicable and explicitly mentioned, p-values were adjusted for a False Discovery Rate of 0.05
19 according to Benjamini & Hochberg, 1995 procedure as implemented in [Groope, 2024].

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

References

- 1
- 2 [Adams et al., 2023] Adams, J.L., Kangarloo, T., Tracey, B. *et al.* Using a smartwatch and
3 smartphone to assess early Parkinson's disease in the WATCH-PD study. *NPI Parkinsons Dis.* **9**,
4 64 (2023). <https://doi.org/10.1038/s41531-023-00497-x>
- 5 [Armstrong & Okun, 2020] Armstrong MJ, Okun MS. Diagnosis and Treatment of Parkinson
6 Disease: A Review. *JAMA.* 2020;323(6):548–560. doi:10.1001/jama.2019.22360
- 7 [Bloem et al., 2019] B. R. Bloem, W. J. Marks Jr., A. L. Silva de Lima, M. L. Kuijf, T. van Laar, B. P.
8 F. Jacobs, M. M. Verbeek, R. C. Helmich, B. P. van de Warrenburg, L. J. W. Evers, J. int Hout, T.
9 van de Zande, T. M. Snyder, R. Kapur & M. J. Meinders. The Personalized Parkinson Project:
10 examining disease progression through broad biomarkers in early Parkinson's disease. *BMC*
11 *Neurol* 19, 160 (2019). <https://doi.org/10.1186/s12883-019-1394-3>
- 12 [Bloem et al., 2021] Bastiaan R Bloem, Michael S Okun, Christine Klein. Parkinson's disease.
13 *The Lancet*, Volume 397, Issue 10291, 2021, Pages 2284-2303. DOI: 10.1016/S0140-
14 6736(21)00218-X.
- 15 [Del Din et al., 2016] Del Din, S., Godfrey, A., Mazzà, C., Lord, S. and Rochester, L. (2016), Free-
16 living monitoring of Parkinson's disease: Lessons from the field. *Mov Disord.*, 31: 1293-1313.
17 <https://doi.org/10.1002/mds.26718>
- 18 [Dirkx et al., 2019] Dirkx MF, Zach H, van Nuland A, Bloem BR, Toni I, Helmich RC. Cerebral
19 differences between dopamine-resistant and dopamine-responsive Parkinson's tremor. *Brain.*
20 2019;142(10):3144-3157. doi:10.1093/brain/awz261
- 21 [Dirkx et al., 2020] Dirkx MF, Zach H, van Nuland AJ, Bloem BR, Toni I, Helmich RC. Cognitive
22 load amplifies Parkinson's tremor through excitatory network influences onto the
23 thalamus. *Brain.* 2020;143(5):1498-1511. doi:10.1093/brain/awaa083
- 24 [Dirkx et al., 2023] Dirkx MF, Shine JM, Helmich RC. Integrative Brain States Facilitate the
25 Expression of Parkinson's Tremor. *Mov Disord.* 2023;38(9):1615-1624. doi:10.1002/mds.29506
- 26 [Dorsey et al., 2018] Dorsey ER, Sherer T, Okun MS, Bloem BR. The Emerging Evidence of the
27 Parkinson Pandemic. *J Parkinsons Dis.* 2018;8(s1): S3-S8. doi: 10.3233/JPD-181474.
- 28 [Elble et al., 2006] R. J. Elble, S. L. Pullman, J. Y. Matsumoto, J. Raethjen, G. Deuschl, R. Tintner,
29 Tremor amplitude is logarithmically related to 4- and 5-point tremor rating scales. *Brain*, v.
30 129(10), pp. 2660–2666, 2006. DOI: [10.1093/brain/awl190](https://doi.org/10.1093/brain/awl190).
- 31 [Evers et al., 2020] L. J. Evers, Y. P. Raykov, J. H. Krijthe, A. L. Silva de Lima, R. Badawy, K. Claes,
32 T. M. Heskes, M. A. Little, M. J. Meinders, B. R. Bloem. Real-Life Gait Performance as a Digital
33 Biomarker for Motor Fluctuations: The Parkinson@Home Validation Study. *Journal of Medical*
34 *Internet Research*, v. 22 (10): e19068, 2020. DOI: 10.2196/19068
- 35 [Farzanehfar et al., 2022] P. Farzanehfar, H. Woodrow, M. Horne. Sensor Measurements Can
36 Characterize Fluctuations and Wearing Off in Parkinson's Disease and Guide Therapy to

1 Improve Motor, Non-motor and Quality of Life Scores. *Front. Aging Neurosci.* 14:852992,
2 2022. DOI: 10.3389/fnagi.2022.852992

3 [Ferreira et al., 2015] Ferreira, J.J., Godinho, C., Santos, A.T. et al. Quantitative home-based
4 assessment of Parkinson's symptoms: The SENSE-PARK feasibility and usability study. *BMC*
5 *Neurol* 15, 89 (2015). <https://doi.org/10.1186/s12883-015-0343-z>

6 [Virbel-Fleischman et al., 2023] Virbel-Fleischman, C., Mousin, F., Liu, S. et al. Symptoms
7 assessment and decision to treat patients with advanced Parkinson's disease based on
8 wearables data. *npj Parkinsons Dis.* 9, 45 (2023). <https://doi.org/10.1038/s41531-023-00489->
9 x

10 [Groope, 2024] David Groope (2024). `fdr_bh` ([https://www.mathworks.com/matlabcentral/](https://www.mathworks.com/matlabcentral/fileexchange/27418-fdr_bh)
11 [fileexchange/27418-fdr_bh](https://www.mathworks.com/matlabcentral/fileexchange/27418-fdr_bh)), MATLAB Central File Exchange. Retrieved January 22, 2024.

12 [Hammerla et al., 2013] N. Y. Hammerla, R. Kirkham, P. Andras, T. Ploetz. On preserving
13 statistical characteristics of accelerometry data using their empirical cumulative distribution.
14 In: *Proceedings of the 2013 international symposium on wearable computers* pp. 65–68, 2013.
15 DOI: 10.1145/2493988.2494353

16 [Heida et al., 2013] T. Heida, E. C. Wentink, E. Marani. Power spectral density analysis of
17 physiological, rest and action tremor in Parkinson's disease patients treated with deep brain
18 stimulation. *J NeuroEngineering Rehabil*, v. 10, pp. 70 (2013). DOI: 10.1186/1743-0003-10-70

19 [Heijmans et al., 2019] M. Heijmans, J. G. V. Habets, C. Herff, *et al.* Monitoring Parkinson's
20 disease symptoms during daily life: a feasibility study. *npj Parkinsons Dis.* 5, 21, 2019. DOI:
21 10.1038/s41531-019-0093-5

22 [Hssayeni et al., 2019] Hssayeni MD, Jimenez-Shahed J, Burack MA, Ghoraani B. Wearable
23 Sensors for Estimation of Parkinsonian Tremor Severity during Free Body Movements. *Sensors*
24 (Basel). 2019 Sep 28;19(19):4215. doi: 10.3390/s19194215.

25 [Mahadevan et al., 2020] N. Mahadevan, C. Demanuele, H. Zhang, D. Volfson, B. Ho, M. K. Erb,
26 S. Patel. Development of digital biomarkers for resting tremor and bradykinesia using a wrist-
27 worn wearable device. *NPJ Digit Med.* v. 15(3:5), 2020 DOI: 10.1038/s41746-019-0217-7.

28 [Kim et al., 2009] S. J. Kim, K. Koh, S. Boyd, D. Gorinevsky. L1 Trend Filtering. *SIAM Review*, v.
29 51(2): 339-360, 2009. DOI: 10.1137/070690274

30 [Lipsmeier et al., 2022] Lipsmeier, F., Taylor, K.I., Postuma, R.B. *et al.* Reliability and validity of
31 the Roche PD Mobile Application for remote monitoring of early Parkinson's disease. *Sci*
32 *Rep* 12, 12081 (2022). <https://doi.org/10.1038/s41598-022-15874-4>

33 [Luft et al., 2019] F. Luft, S. Sharifi, W. Mugge, A. C. Schouten, L. J. Bour, A-F Rootselaar, P. H.
34 Veltinik & T. Heida. A Power Spectral Density-Based Method to Detect Tremor and Tremor
35 Intermittency in Movement Disorders. *Sensors*, v. 19: 4301, 2019. DOI: 10.3390/s19194301

36 [Ossig et al., 2016] C. Ossig, F. Gandor, M. Fauser, C. Bosredon, L. Churilov, H. Reichmann, M.
37 K. Horne, G. Ebersbach, A. Storch. Correlation of Quantitative Motor State Assessment Using

1 a Kinetograph and Patient Diaries in Advanced PD: Data from an Observational Study. PLoS
2 ONE 11(8):e0161559, 2016. DOI:10.1371/journal.pone.0161559

3 [Ou et al., 2021] Ou Z, Pan J, Tang S, Duan D, Yu D, Nong H, Wang Z. Global Trends in the
4 Incidence, Prevalence, and Years Lived With Disability of Parkinson's Disease in 204
5 Countries/Territories From 1990 to 2019. *Front Public Health*. 2021 Dec 7;9:776847. doi:
6 10.3389/fpubh.2021.776847.

7 [Oyama et al., 2023] G. Oyama, M. Burq, T. Hatano, et al. Analytical and clinical validity of
8 wearable, multi-sensor technology for assessment of motor function in patients with
9 Parkinson's disease in Japan. *Sci Rep* 13, 3600, 2023. DOI: 10.1038/s41598-023-29382-6

10 [Patel et al., 2009] S. Patel; K. Lorincz; R. Hughes; N. Huggins; J. Growdon; D. Standaert; M.
11 Akay; J. Dy; M. Welsh, P. Bonato. Monitoring Motor Fluctuations in Patients With Parkinson's
12 Disease Using Wearable Sensors. *IEEE Transactions on Information Technology in Biomedicine*,
13 vol. 13 (6), pp. 864-873, 2009. DOI: 10.1109/TITB.2009.2033471.

14 [Perepezko et al., 2020] K. Perepezko, J. Hinkle, K. Mills, G. Pontone. Frequency of Wearing-
15 Off Symptoms in Parkinson's disease Fluctuators: An on/off evaluation [abstract]. *Mov*
16 *Disord*. 2020; 35 (suppl 1). [https://www.mdsabstracts.org/abstract/frequency-of-wearing-](https://www.mdsabstracts.org/abstract/frequency-of-wearing-off-symptoms-in-parkinsons-disease-fluctuators-an-on-off-evaluation/)
17 [off-symptoms-in-parkinsons-disease-fluctuators-an-on-off-evaluation/](https://www.mdsabstracts.org/abstract/frequency-of-wearing-off-symptoms-in-parkinsons-disease-fluctuators-an-on-off-evaluation/). Accessed February
18 28, 2025.

19 [Pulliam et al., 2018] C. L. Pulliam, D. A. Heldman, E. B. Brokaw, T. O. Mera, Z. K. Mari, M. A.
20 Burack. Continuous Assessment of Levodopa Response in Parkinson's Disease Using Wearable
21 Motion Sensors. *IEEE Trans Biomed Eng*. v. 65(1):159-164, 2018 DOI:
22 10.1109/TBME.2017.2697764.

23 [Maetzler et al., 2013] Maetzler W, Domingos J, Srulijes K, Ferreira JJ, Bloem BR. Quantitative
24 wearable sensors for objective assessment of Parkinson's disease. *Mov Disord*. 2013
25 Oct;28(12):1628-37. doi: 10.1002/mds.25628.

26 [McGurrin et al., 2021] P. McGurrin, J. Mcnames, T. Wu, M. Hallet & D. Haubenberger.
27 Quantifying Tremor in Essential Tremor Using Inertial Sensor-Validation of an Algorithm. *IEEE*
28 *Journal of Translational Engineering in Health and Medicine*, v. 9:2700110, 2021. DOI:
29 10.1109/JTEHM.2020.3032924

30 [Moreau et al., 2023] Moreau, C., Rouaud, T., Grabli, D. *et al*. Overview on wearable sensors
31 for the management of Parkinson's disease. *npj Parkinsons Dis*. 9, 153 (2023). DOI:
32 10.1038/s41531-023-00585-y

33 [Rovini et al., 2017] Rovini E, Maremmani C, Cavallo F. How Wearable Sensors Can Support
34 Parkinson's Disease Diagnosis and Treatment: A Systematic Review. *Front Neurosci*. 2017 Oct
35 6;11:555. doi: 10.3389/fnins.2017.00555.

36 [San-Segundo et al., 2016] R. San-Segundo, J. M. Montero, R. Barra-Chicote, F. Fernández, J.
37 M. Pardo. Feature extraction from smartphone inertial signals for human activity
38 segmentation. *Signal Processing*, v. 120, pp. 359-372, 2016. DOI:
39 10.1016/j.sigpro.2015.09.029.

1 [San-Segundo et al., 2020] R. San-Segundo, A. Zhang, A. Cebulla, S. Panev, G. Tabor, K.
2 Stebbins, R. E. Massa, A. Whitford, F. de la Torre, J. Hodgins. Parkinson's Disease Tremor
3 Detection in the Wild Using Wearable Accelerometers. *Sensors* 2020, 20, 5817. DOI:
4 10.3390/s20205817

5 [Sabatini, 2011] Sabatini AM. Kalman-filter-based orientation determination using
6 inertial/magnetic sensors: observability analysis and performance evaluation. *Sensors (Basel)*.
7 2011;11(10):9182-206. doi: 10.3390/s111009182.

8 [Schafer & Markel, 1979] R. W. Schafer, J. D. Markel. *Speech Analysis*. John Wiley & Sons Inc,
9 1979

10 [Sica et al., 2021] Sica M, Tedesco S, Crowe C, Kenny L, Moore K, et al. (2021) Continuous home
11 monitoring of Parkinson's disease using inertial sensors: A systematic review. *PLOS ONE* 16(2):
12 e0246528. <https://doi.org/10.1371/journal.pone.0246528>

13 [Sigcha et al., 2021] L. Sigcha, I. Pavón, N. Costa, S. Costa, M. Gago, P. Arezes, J. M. López, G.
14 Arcas. Automatic Resting Tremor Assessment in Parkinson's Disease Using Smartwatches and
15 Multitask Convolutional Neural Networks. *Sensors (Basel)*. 2021 Jan 4;21(1):291. DOI:
16 10.3390/s21010291.

17 [Stacy et al., 2005] Stacy M, Bowron A, Guttman M, Hauser R, Hughes K, Larsen JP, LeWitt P,
18 Oertel W, Quinn N, Sethi K, Stocchi F. Identification of motor and nonmotor wearing-off in
19 Parkinson's disease: Comparison of a patient questionnaire versus a clinician assessment.
20 *Movement Disorders*, 2005, v. 20(6): 726-733. DOI: <https://doi.org/10.1002/mds.20383>

21 [Swinnen et al., 2025] Swinnen BEKS, Frequin HL, Wiggerts Y, Espay AJ, Beudel M, de Bie RMA.
22 Tremor Is Highly Responsive to Levodopa in Advanced Parkinson's Disease. *Mov Disord Clin*
23 *Pract*. 2025;12(1):76-81. doi:10.1002/mdc3.14262

24 [Tolosa et al., 2021] Tolosa E, Garrido A, Scholz SW, Poewe W. Challenges in the diagnosis of
25 Parkinson's disease. *Lancet Neurol*. 2021 May;20(5):385-397. doi: 10.1016/S1474-
26 4422(21)00030-2.

27 Tzallas AT, Tsipouras MG, Rigas G, Tsalikakis DG, Karvounis EC, Chondrogiorgi M, Psomadellis
28 F, Cancela J, Pastorino M, Waldmeyer MT, Konitsiotis S, Fotiadis DI. PERFORM: a system for
29 monitoring, assessment and management of patients with Parkinson's disease. *Sensors*
30 (Basel). 2014 Nov 11;14(11):21329-57. doi: 10.3390/s141121329.

31 [Varghese et al., 2024] Varghese, J., Brenner, A., Fujarski, M. et al. Machine Learning in the
32 Parkinson's disease smartwatch (PADS) dataset. *npj Parkinsons Dis*. 10, 9 (2024).
33 <https://doi.org/10.1038/s41531-023-00625-7>

34 [Zach et al., 2015] Zach H, Dirx M, Bloem BR, Helmich RC. The Clinical Evaluation of Parkinson's
35 Tremor. *J Parkinsons Dis*. 2015;5(3):471-474. doi:10.3233/JPD-150650

36 [Zach et al., 2017] Zach, H., Dirx, M.F., Pasma, J.W., Bloem, B.R. and Helmich, R.C. (2017),
37 Cognitive Stress Reduces the Effect of Levodopa on Parkinson's Resting Tremor. *CNS Neurosci*
38 *Ther*, 23: 209-215. <https://doi.org/10.1111/cns.12670>

1 [Zach et al., 2020] Zach H, Dirkx MF, Roth D, Pasman JW, Bloem BR, Helmich RC. Dopamine-
2 responsive and dopamine-resistant resting tremor in Parkinson disease. *Neurology*. 2020 Sep
3 15;95(11):e1461-e1470. doi: 10.1212/WNL.0000000000010316.

4 [Zhao & Wang, 2013] X. Zhao and D. Wang, "Analyzing noise robustness of MFCC and GFCC
5 features in speaker identification," 2013 IEEE International Conference on Acoustics, Speech
6 and Signal Processing, Vancouver, BC, Canada, 2013, pp. 7204-7208, DOI:
7 10.1109/ICASSP.2013.6639061.

8

9

10

11

12

13

14

15

16

17

Supplementary Material

18

Methods

19

Supervised Tremor and arm activity detection

20

21

22

23

24

25

26

27

28

29

30

Data-Processing: Tri-axial acceleration and angular velocity data were acquired by Physilog®
4 (Gait Up S.A., Lausanne, Switzerland) sensor displaced at both wrists, ankles, and lower
back. Data from the most tremor-affected upper limb was considered in this analysis. Data
were acquired at 200 Hz and further downsampled to 100 Hz. Signals were sliced into 4 s
windows with a 2 s overlap. Gravity effects on the acceleration recordings were corrected
according to the L1 trend filtering [Kim et al., 2009]. Details concerning experimental setup
for signal acquisition can be found in [Evers et al., 2020]. Feature extraction was performed
considering: (1) Power spectral density (PSD) attributes; (2) Mel-frequency cepstral
coefficients (MFCCs); (3) Data-range-based features obtained by the empirical cumulative
distributions functions (ECDFs); Root mean square (RMS) attributes. Evaluation of these
attributes is detailed in the Feature Extraction section in this supplementary material.

31

32

33

34

35

36

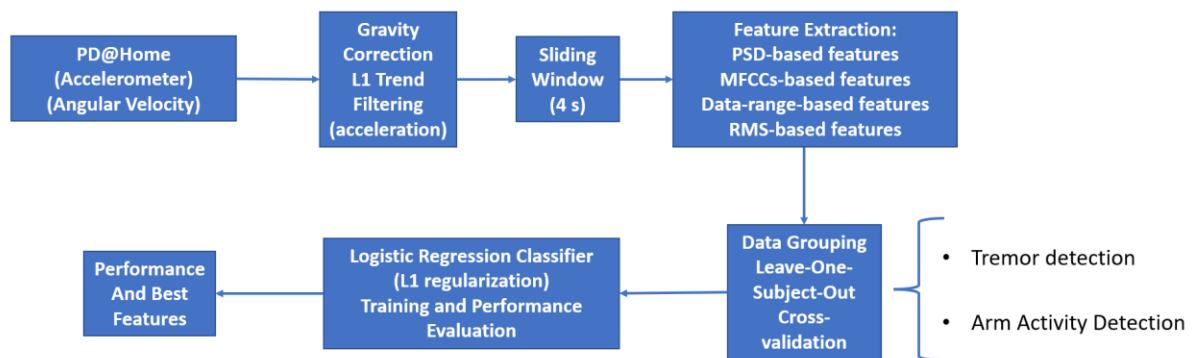
37

Based on the video annotations provided, window labels were defined independently for
activity (walking, sitting, standing, etc) and tremor (see details in [Evers et al., 2020]). Tremor
label under significant arm activity is defined in terms of a binary variable (occurrence or not),
while under rest, information concerning severity is also provided (No Tremor, Mild Tremor,
Moderate Tremor, Severe Tremor). Tremor episodes were detected based on the following
binary supervised classification scenario: class NO TREMOR (class 1 - negative) refers to the
condition of no tremor; class TREMOR (class 2 - positive) is characterized by tremor

1 independently of severity. A second classifier is implemented to detect significant arm-activity
 2 (i.e., independently from the specific activity being performed). In this case, class 1 - NO ARM
 3 ACTIVITY - is the negative one, and class 2 – ARM ACTIVITY is the positive one. For both
 4 classification scenarios, 4 s windows with at least 1 s of tremor are labelled as TREMOR, and
 5 windows with at least 0.5 s of arm activity are labelled as ARM ACTIVITY.

6 Detection performance was evaluated considering seven PD patients with tremor in a leave-
 7 one subject out (LOSO) crossvalidation scheme using a logistic regression classifier under least
 8 absolute shrinkage and selection operator (LASSO) regularization for each classification task
 9 (tremor and arm activity). Performance was characterized through the receiver operating
 10 characteristics curve (ROC). The signal processing pipeline is schematically presented in Figure
 11 **S1**.

12



13

14 Figure **S1**: Tremor detection and arm activity signal processing pipeline. Tri-axial acceleration and
 15 angular velocity were pre-processed and sliced into windows. Different set of features were extracted,
 16 and the performance was evaluated through a leave one subject out cross-validation scheme using a
 17 logistic regression classifier for tremor and arm activity detection.

18

19 Feature Extraction

20 **PSD-based features**: power spectral density coefficients (PSD) outline a classical approach for
 21 investigating and characterizing tremor and general activity events in wearables sensors [Luft
 22 et al., 2019]. In this work, the PSD coefficients were obtained for each axis ($P_x(f)$, $P_y(f)$, $P_z(f)$)
 23 and collapsed (summed) into a single vector ($P_s(f) = P_x(f) + P_y(f) + P_z(f)$) for acceleration and
 24 angular velocity. This provides a lower dimensional and orientation independent feature
 25 structure [McGurrian et al., 2021] for each inertial measurement. For each type of recording,
 26 seven PSD-based attributes were estimated for 3 different bands: (1) **BandPower** – area
 27 under the PSD functions within the respective bandwidth; (2) **BandPowerRatio** – BandPower
 28 normalized by the sum of coefficients up to 15 Hz; (3) **PeakFrequency** – dominant frequency
 29 within the bandwidth; (4) **DominantPower** – area under the dominant frequency constrained
 30 to half-decay range within the bandwidth; (5) **DominantPowerRatio** – dominant power
 31 normalized by the BandPower; (6) **FrequencyWidth** – length of the frequency interval around
 32 the dominant peak for half-decay within the bandwidth; (7) **PeakHeight** – difference between
 33 the frequency peak and the baseline within the bandwidth. The bandwidths considered were:

1 [0.4 – 2 Hz] typically associated with gait activity; [3.5 - 8 Hz] typically associated with rest
2 tremor activity; typically [8 – 12 Hz] associated with postural and kinetic tremor activity. This
3 spectral segregation aims an efficient feature set description considering the essential activity
4 detection phenomena to be considered.

5 **Mel-frequency cepstral coefficients (MFCCs):** MFCCS define a small set of features that
6 concisely describe the overall shape of the spectral envelope [Schafer & Markel, 1979]. The
7 approach used here starts by obtaining the Short Time Fourier Transform of each axis
8 (Hamming sub-windows of 2 s with 80% overlap), followed by their sum for obtaining an
9 orientation-independent feature structure. The resulting spectrogram is then filtered by 15
10 triangular filters distributed in the range [0 - 30 Hz] with filter edges defined by the Mel-
11 adjusted scale for inertial signals:

$$12 \quad \text{Mel}(f) = 64.875 \cdot \log_{10} \left(1 + \frac{f}{17.5} \right) \quad (1).$$

13 The output of the bank filter is then non-linear ratified by a log function and cosine discrete
14 transformed (DCT-II) providing 12 coefficients. Essentially, the MFCCs focuses the analysis on
15 lower frequencies, using the Mel scale for obtaining a sparser distribution of the filters in the
16 upper part of the spectral domain. The scale mimics the non-linear human sound perception
17 which justifies its adaption for working within the range of inertial signals as shown in [San-
18 Segundo, 2016].

19 **Data-range features:** Data-range features were evaluated based on the empirical cumulative
20 distribution functions (ECDFs) as introduced in the context of accelerometry by [Hammerla et
21 al., 2013]. ECDFs corresponds to a set of concatenated quantiles of the data distribution,
22 which claims to define a more robust statistics, since different movement activities can imply
23 multimodal distributions of difficult characterization by standard statistical moments (e.g., by
24 imprecise definition of data range or histogram bins). For the evaluation employed here, the
25 data is firstly filtered in different bandwidths: [0.4 – 2 Hz] gait range; [3.5 – 8Hz] rest tremor
26 range; [8 – 12 Hz] postural/kinetic tremor range; [12- 20 Hz] higher frequency activity. For
27 each frequency range, tri-axial data was collapsed by evaluating the magnitude vector:

$$28 \quad \text{Mag}(n) = \sqrt{s_x^2(n) + s_y^2(n) + s_z^2(n)} \quad (2),$$

29 in which $s_i(n)$ denotes the n -th sample of the signal value at axis i . The ECDF is then obtained
30 from $\text{Mag}(n)$ by means of 10 concatenated quantiles equally distributed in the range [0 1].
31 The feature set also includes the distribution mean, leading to 11 attributes per bandwidth.

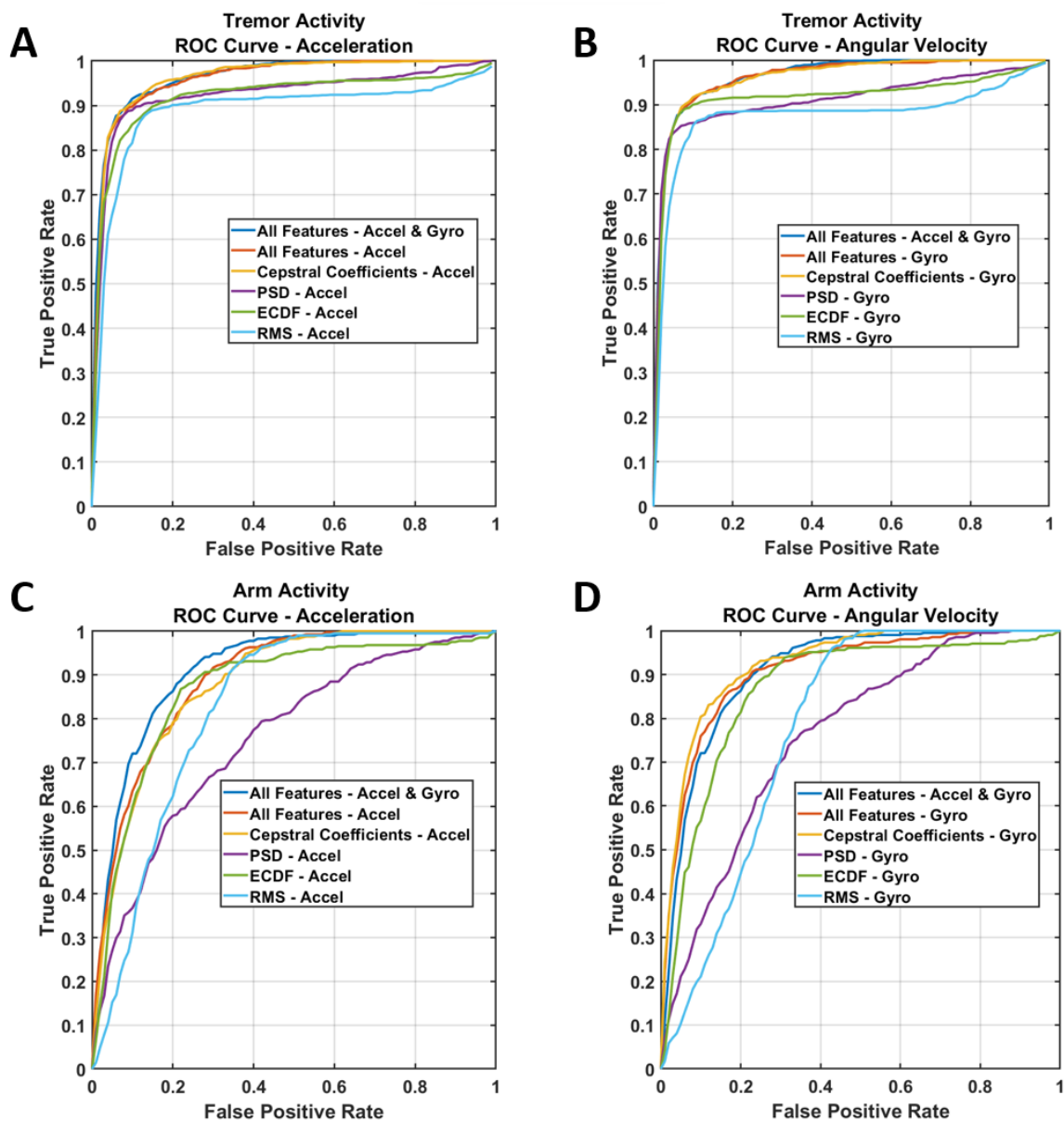
32 **Root Mean Square (RMS)-based features:** the RMS value of inertial signals can also outline a
33 low-cost measure for characterizing tremor or other movement activities [Patel et al., 2009].
34 For each bandwidth considered in the data-range features, the RMS value of the magnitude
35 vector was obtained, leading to 4 attributes.

36 **Data Partition and Classification Scenarios:** A leave-one-subject-out (LOSO) cross-validation
37 scheme and a logistic regression classifier under least absolute shrinkage and selection
38 operator (LASSO) regularization were used to evaluate classification performance. All
39 biomarkers' candidates were z-scored normalized based on the respective feature average

1 and standard deviation estimated from the training set. Independent classifiers were
2 obtained for tremor and arm activity detection. LASSO penalty was set to constrain the
3 maximum number of features to 20 in each classification scenario, which aims to explore the
4 underlying information of the feature space, while preserving a reasonable low dimensional
5 representative feature set. The choice of the logistic regression classifier is motivated by its
6 simplicity, interpretability and widespread use in the binary classification tasks. Detection
7 performance is characterized through the receiver operating characteristics (ROC) curve and
8 the area under it (AUC), which were obtained for each type of inertial measurement
9 (acceleration and angular velocity), for each type of feature and for the combination of all
10 attributes/type of measurement.

11 **Supervised Detection Performance:** Figure **S2** shows the ROC curves for tremor (Figs. **S2(A)**
12 and **S2(B)**) and arm activity (Figs. **S2(C)** and **S2(D)**) for each set of features extracted and their
13 combinations. Feature set combination includes all features extracted under acceleration
14 measurements (All features - Accel) and all features extracted for both acceleration and
15 angular velocity (All features – Accel & Gyro). From the ROC curves, it is possible to observe
16 that the cepstral coefficients (MFCCs) have the best performance among the individual set of
17 features, providing detection characteristics identical to all features combined (Figs. **S2(A)**,
18 **S2(B)**, and **S2(D)**). A small performance difference is only observed for arm activity detection
19 using acceleration recordings (Fig. **S2(C)**). The analysis emphasizes crucial aspects for
20 optimizing the detection approach: (1) there is redundant information among the features
21 sets; (2) there is redundant information between acceleration and angular velocity.

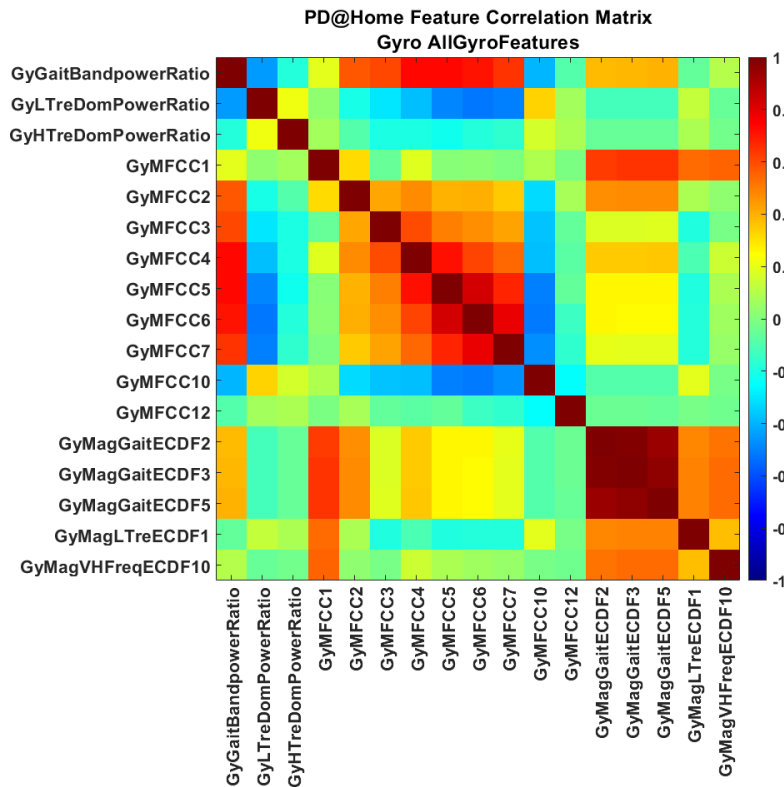
22



1

2 Figure S2: (A) Receiver operating characteristics (ROC) curve for the tremor detection task considering
 3 all individual set of attributes extracted and their combination among acceleration recordings (All
 4 features – Accel) and all features independent of inertial measurement type (All Features – Accel &
 5 Gyro); (B) Analogous performance characterization for the angular velocity measurement; (C) ROC
 6 curve for the arm activity detection task considering acceleration, and (D) angular velocity.

7 Indeed, Figure S3 shows the correlation between the best gyroscope features, i.e., attributes
 8 selected at least 4 times in the LOSO cross-validation scheme by the logistic regression
 9 classifier under LASSO regularization. Redundancy across feature sets is exemplified by the
 10 correlation between the first MFCC coefficient with ECDF features, or with MFCCs in general
 11 with PSD-based features (e.g., GyGaitBandpowerRatio, GyLTreDomPowerRatio). This
 12 redundant behaviour suggests that the MFCCs set, which was predominant in the best feature
 13 set selected, can efficiently summarize the information content among all the feature sets
 14 used.

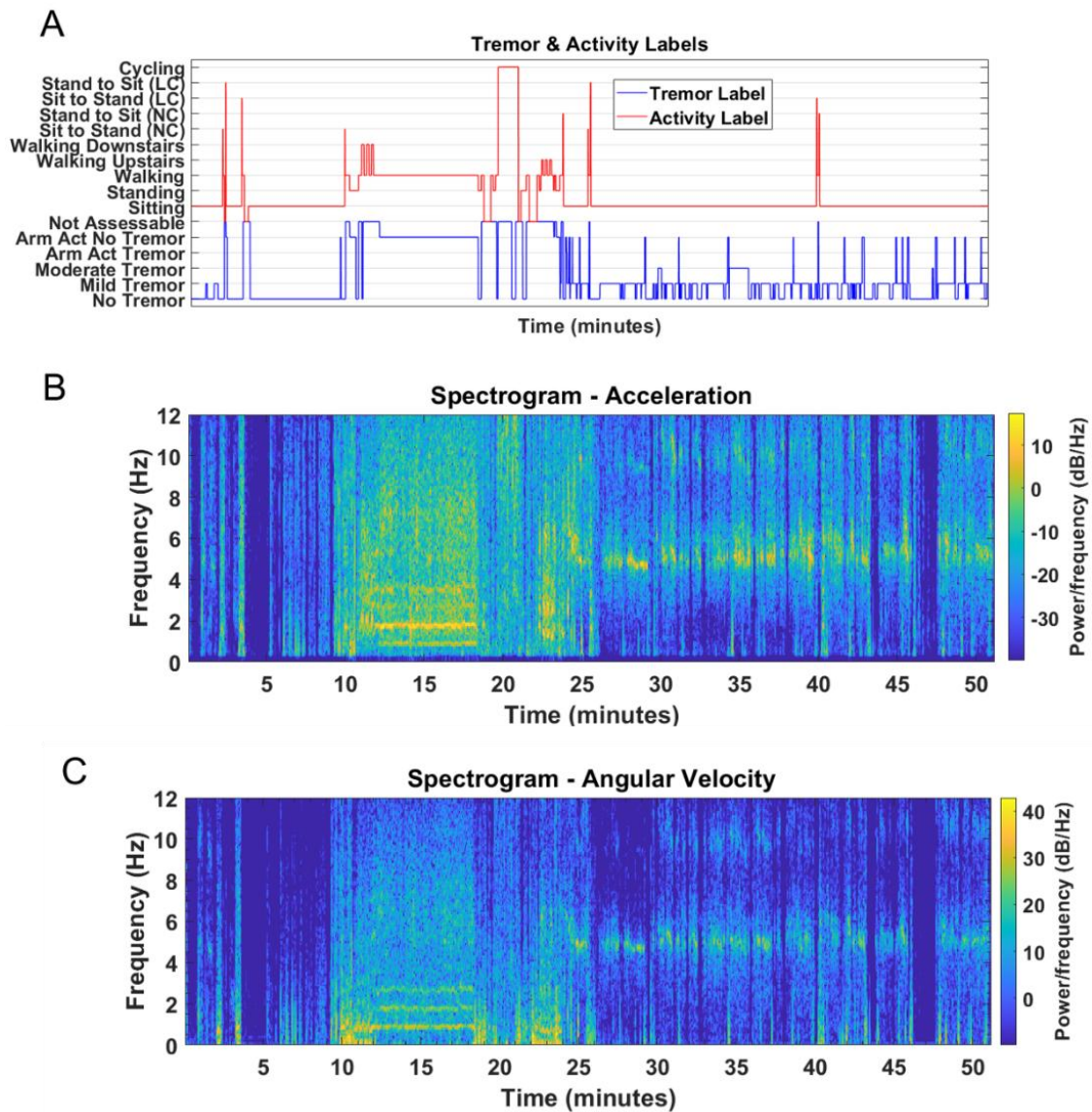


1

2 Figure **S3**: Correlation matrix among the best features selected by logistic regression
 3 classification for tremor detection using gyroscope measurements. **GyGaitBandpowerRatio** –
 4 gyroscope bandpower ratio in gait frequency range ([0.4 - 2 Hz]), **GyLTreDomPowerRatio** –
 5 gyroscope dominant power ratio in the range [3.5 – 8 Hz], **GyHTreDomPowerRatio** –
 6 gyroscope dominant power ratio in the range [8 – 12 Hz], **GyMFCC(*i*)** – *i*-th MFCC coefficient
 7 from gyroscope recordings, **GyMagGaitECDF(*k*)** – *k*-th ECDF coefficient for the measurement
 8 magnitude in the range [0.4 – 2 Hz], **GyMagLTreECDF(*k*)** – *k*-th coefficient for the magnitude
 9 in the range [3.5 – 8 Hz], **GyMagVHFreqECDF(*k*)** – *k*-th ECDF coefficient for the measurement
 10 magnitude in the range [12-20 Hz].

11

12 Fig. **S4(A)** shows a typical labeled time course for patient’s tremor and the respective
 13 ongoing activity as labelled in video recording. Spectrograms for acceleration (Fig. **S4(B)**) and
 14 angular velocity (Fig. **S4(C)**) show similar time-frequency structures and redundant spectral
 15 information between the inertial measurements. The spectrograms also indicate different
 16 harmonic structures for at least two essential daily activities: walking (e.g., from the 12th to
 17 18th minute) and tremor (e.g., such as intermittently appearing after the 25th minute). Indeed,
 18 walking activity introduces harmonics within the typical tremor band (3.5-12 Hz), contributing
 19 as hard confounding effect if tremor detection would be performed just based in the typical
 20 range-based bandpower estimates. In this case, the “spectral landscape” captured by the
 21 MFCCs seems particularly efficient to deal with the different harmonic structures observed in
 22 walking and tremor activities and may explain why they were the best feature set among the
 23 attributes extracted (Figure **S2**).



1

2 Figure S4: (A) Labelled tremor and activity time course for a typical PD@Home recorded
 3 section. “LC” and “NC” denote “low chair” and “normal chair”, respectively. “Act” denotes
 4 activity. Only the tremor label was used to obtain the tremor and arm activity classifiers; (B)
 5 Summed spectrograms for the 3-axis acceleration and (C) for 3-axis angular velocity.

6

7

8 The results shown in Figures S2, S3 and S4 leads to key-conclusion to optimize the
 9 classification approach: the evaluation of the MFCCs is enough to attain best detection results
 10 both for tremor and arm activity detection when gyroscope measurements are considered.
 11 This avoids dealing with computationally expensive gravity correction techniques in
 12 accelerometry and outlines a small set of frequency-based descriptors capable for capturing
 13 different complex harmonic structures without *a priori* assumptions about specific activity
 14 bandwidths. Table S1 shows the final LOSO performance considering MFCCs extracted from
 15 angular velocity. For tremor detection, the logistic regression threshold was chosen to have a

1 specificity of 95 % aiming at controlling the negative rate detection, given the expected class
2 balance in long-term evaluation. For tremor detection, all twelve MFCCs features were used,
3 except coefficients 9 and 11. For arm activity detection, the MFCCs 1, 3, 4, 5, and 6 were used.

4 Table **S1**: Mean \pm standard deviation for sensitivity, specificity and area under the ROC curve
5 for tremor and arm activity detection based on MFCCs extracted from gyroscope data.

6

CLASSIFICATION TASK	SENSITIVITY	SPECIFICITY	AUC
Tremor detection	0.65 \pm 0.18	0.95 \pm 0.00	0.91 \pm 0.06
Arm-Activity detection	0.74 \pm 0.17	0.78 \pm 0.15	0.88 \pm 0.06

7

8 **Advantages and relation to previous work.**

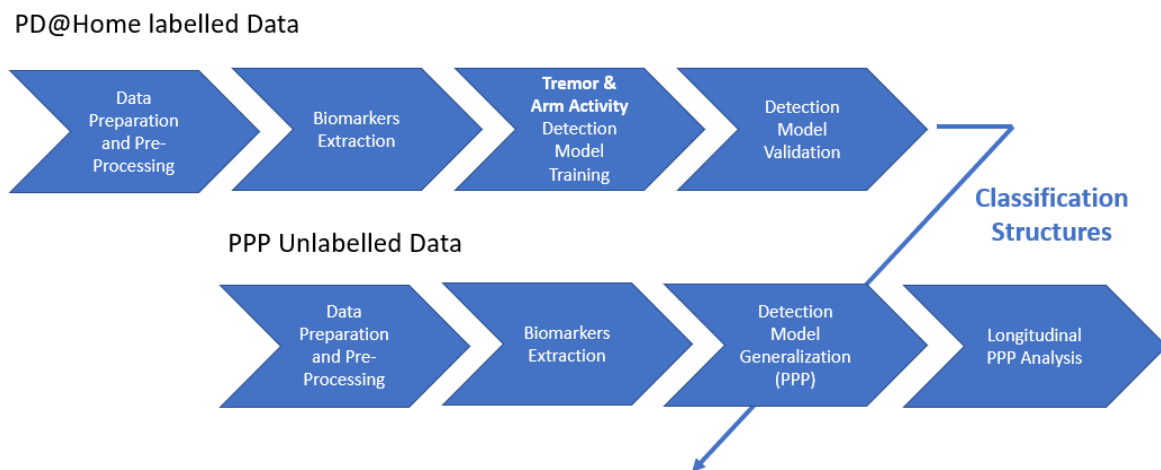
9 The approach stands out by using gyroscope data to avoid computationally expensive gravity
10 effect corrections and a convenient scale-invariant feature set based on the spectral shape
11 and MFCCs, combined with a logistic regression classifier for independently detecting tremor
12 and significant arm activity. In particular, the exclusive use of angular velocity and MFCCs
13 features emerged from a data-driven analysis performed with labelled data, supporting the
14 designing options under a detection performance with AUC > 0.8. It is important to note that
15 only 6.6 % of metrics proposed for free living activity assessments exclusively use gyroscope
16 recordings [Jung et al., 2020], in order to circumvent computationally intensive algorithms
17 (e.g., sensor fusing [Sabatini, 2011] or l1-trend filtering [Kim et al., 2009]) for correcting the
18 gravity effect in the accelerometry: an issue of major relevance for processing data at the
19 scale of hundreds of patients recorded for months or years. The supervised training analysis
20 also allowed to show that the MFCCs can provide a concise and efficient feature set for tremor
21 detection, avoiding the combination of a myriad of redundant attributes. The use of MFCCs in
22 the context of inertial signals was introduced by [San-Segundo et al., 2016] and has also been
23 successfully applied for tremor detection in PD as shown by [San-Segundo et al., 2020; Sigcha
24 et al., 2021]. These features employ an adapted perceptual-based scale to create a bank filter
25 that favours sampling the activity at lower frequency components. The feature set is also
26 scale-invariant to the original signal [Zhao & Wang, 2013], allowing robust tremor detection
27 irrespective of differences in the device sensitivity to movement, which may arise from
28 different sensors or slightly different device placement across days.

29 Heuristics for continuous rest tremor monitoring in free-living conditions with different
30 degrees of complexity have been proposed as reviewed in [Sigcha et al., 2021; Moreau et al.,
31 2023], which implies choosing among different designing options considering single vs. multi-
32 sensor recordings, different types of inertial signals, multi-modal recordings, different feature
33 sets and classification strategies. In general, tremor detection performances with AUC > 0.8
34 [Pulliam et al., 2018; San-Segundo et al., 2020] or accuracy higher than 85% [Tzallas et al.,

1 2012; Rigas et al., 2012; Mahadevan et al., 2020] in the labelled scenario claim to be
2 associated with good performances.

3 Tremor and arm activity detection for the Personalized Parkinson Project

4 Figure S5 shows a schematically representation of the processing pipeline for obtaining
5 tremor and arm activity labels in the Personalized Parkinson Project (PPP) dataset. Logistic
6 regression classifiers for tremor and arm activity are obtained based on labelled data provided
7 by the PD@Home as previously described (upper processing blocks – Figure S5) and applied
8 to the same biomarkers extracted from the PPP dataset (MFCCs evaluated on 4 s windows of
9 gyroscope data without overlap). Raw gyroscope data was cubic interpolated to adjust the
10 original sampling rate to 50 or 100 Hz, aiming at defining a fixed sampling rate as close as
11 possible to the original one. The MFCCs features were z-score normalized considering the
12 average and standard deviation of the attributes obtained for the initial and week 52 (i.e.,
13 around the first and the second clinical assessments) during daytime ([9:00 – 21:00 h]) for
14 patients with tremor constancy score higher than 1. These conditions aim to match the ones
15 observed for training the classifier under the PD@Home dataset, while also considering data
16 acquisition performed by Verily sensors in the PPP scenario.



17

18 Figure S5: Schematical representation for PPP processing steps. Logistic regression classifiers are
19 trained based on labelled data provided by PD@Home and applied to the same biomarkers extracted
20 for the PPP dataset after z-score normalization. PPP daily average in the frequency domain is obtained
21 from circadian wrist-derived measures (probability and severity) aiming at characterizing the rest
22 tremor fluctuation activity.

23 Wrist-based tremor variables processing and PSD estimation.

24 Circadian wrist-based variables were estimated for consecutive 5 minutes windows without
25 overlap as briefly introduced in the Methods Section. Both tremor probability and severity
26 were estimated within the daily-activity interval [6:00 – 23:00 h] and smoothed by a 6 point
27 (30 minutes) moving average filter. Just days with at least 80% of non-missing data within the
28 interval were considered for daily PSD estimation. Missing data was filled by adjusting an
29 autoregressive model using 3 hours of preceding data by means of the “fillgaps” function
30 available from Matlab R2022b library. Each filled daily-cycle for the wrist variables were z-

1 scored in time aiming to obtain comparable daily tremor patterns. Finally, daily power spectral
2 density coefficients were obtained, and the respective frequency coefficients were averaged
3 across days. Just patients with overall tremor mean probability higher than 1.4 % were
4 considered in the analysis. This aims to avoid patients with small spurious tremor patterns due
5 to false positive tremor detections. This tremor threshold was defined based on the 75th-
6 percentile of the mean tremor distribution for patients with UPDRS 3.17 and UPDRS 3.18
7 scores rated as 0 in ON and OFF conditions in both clinical assessments under the same
8 constrains of the medicated group (i.e., taking just levodopa at least 3 times a day and not
9 taking levodopa agonists), defining a non-tremulous cohort of 37 patients.

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

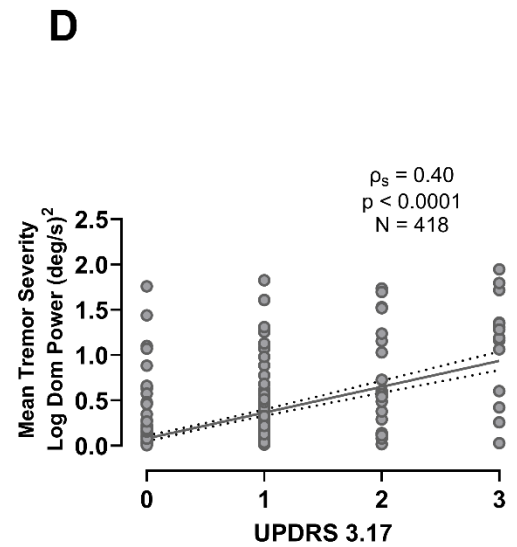
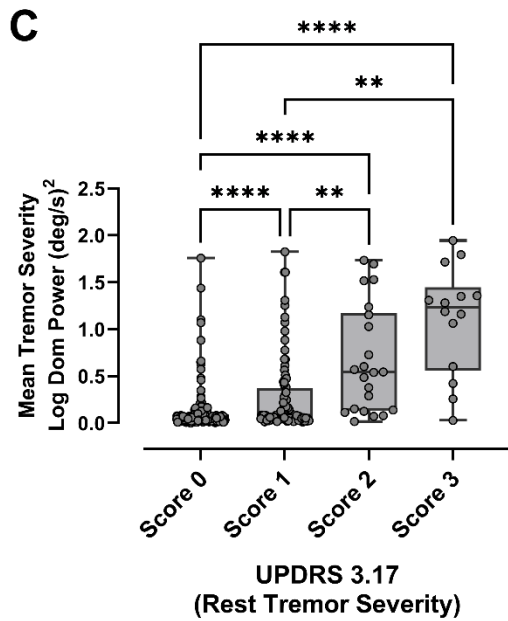
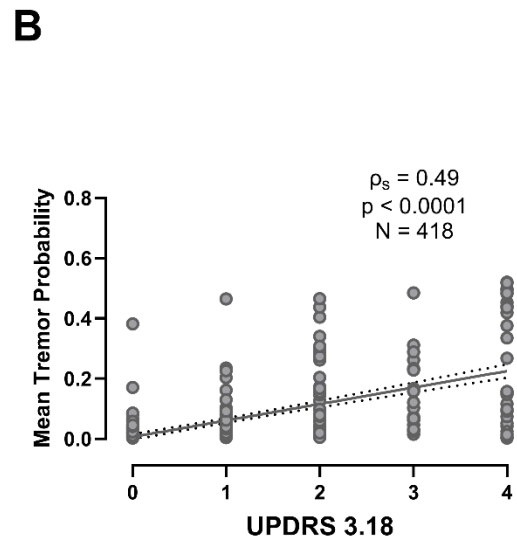
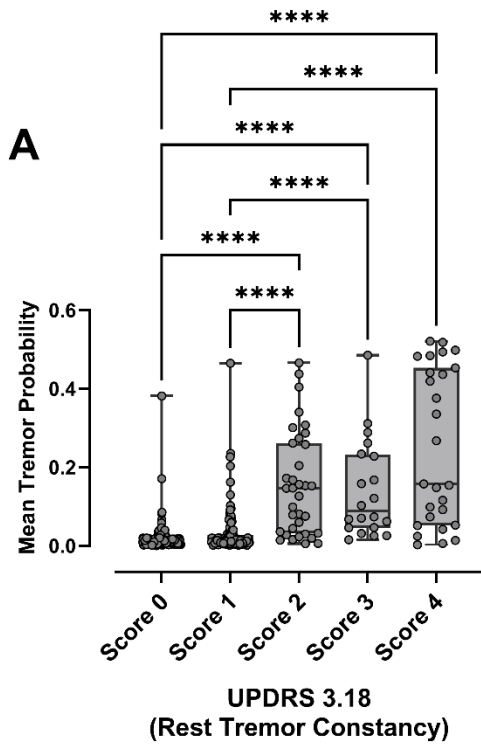
Supplementary Results

28 Tremor detection model and agreement with clinical scores

29 Figure **S6** illustrates the agreement between wrist-derived measures - mean tremor
30 probability and mean tremor severity — and their corresponding clinical scores (rest tremor
31 constancy, UPDRS 3.18, and rest tremor severity, UPDRS 3.17) based on eight weeks of data
32 collected around the second clinical visit. These results are consistent with those observed for
33 the first clinical visit, as reported in Figure **1** (Results Section). Both mean tremor probability

1 and mean tremor severity showed significant differences when stratified by clinical scores (Fig.
2 **S6(A)** and **S6(C)**; Kruskal-Wallis' test with Dunn's post hoc correction). Moreover, wrist-based
3 measures demonstrated significant correlations with their respective clinical scores (Fig. **S6(B)**
4 and **S6(D)**). Specifically, mean tremor probability correlated with UPDRS 3.18 ($\rho_s = 0.49$, $p <$
5 0.0001 , $N = 418$), and mean tremor severity correlated with UPDRS 3.17 ($\rho_s = 0.40$, $p < 0.0001$,
6 $N = 418$).

7



1

2 Figure S6: (A) Boxplots for the wrist-estimated average tremor probability stratified into ascending rest
3 tremor constancy scores (UPDRS 3.18), and (B) the respective dispersion, linear regression, and
4 Spearman correlation value; (C) Boxplots for the wrist-estimated mean tremor severity stratified into
5 ascending rest tremor clinical severity scores (UPDRS 3.17) and (D) the respective dispersion, linear
6 regression, and Spearman correlation value. Data from the second clinical assessment.

7

1 **Tremor fluctuation dynamics – Tremor Severity.**

2 The cohorts analysed were independently defined for tremor constancy and severity based
3 on the response of each respective symptom to levodopa during clinical visits. Table **S2**
4 presents the demographic characteristics of the cohorts, focusing on levodopa responsiveness
5 in relation to tremor severity.

6

7 Table **S2**: Demographic characteristics of the population, categorized by responsiveness to the
8 levodopa challenge based on tremor severity. **NR** represents the levodopa non-responsive
9 cohort, **R** indicates the levodopa responsive cohort, and **NM** refers to the non-medicated
10 cohort. For each variable (except gender), values are reported as averages \pm standard
11 deviation. **LEDD** denotes the levodopa equivalent daily dose; **H&Y** refers to the Hoehn & Yahr
12 scale in the levodopa OFF condition; **UPDRS III** represents the total score of the Unified
13 Parkinson’s Disease Rating Scale - Part III: motor examination in the levodopa OFF condition.

14

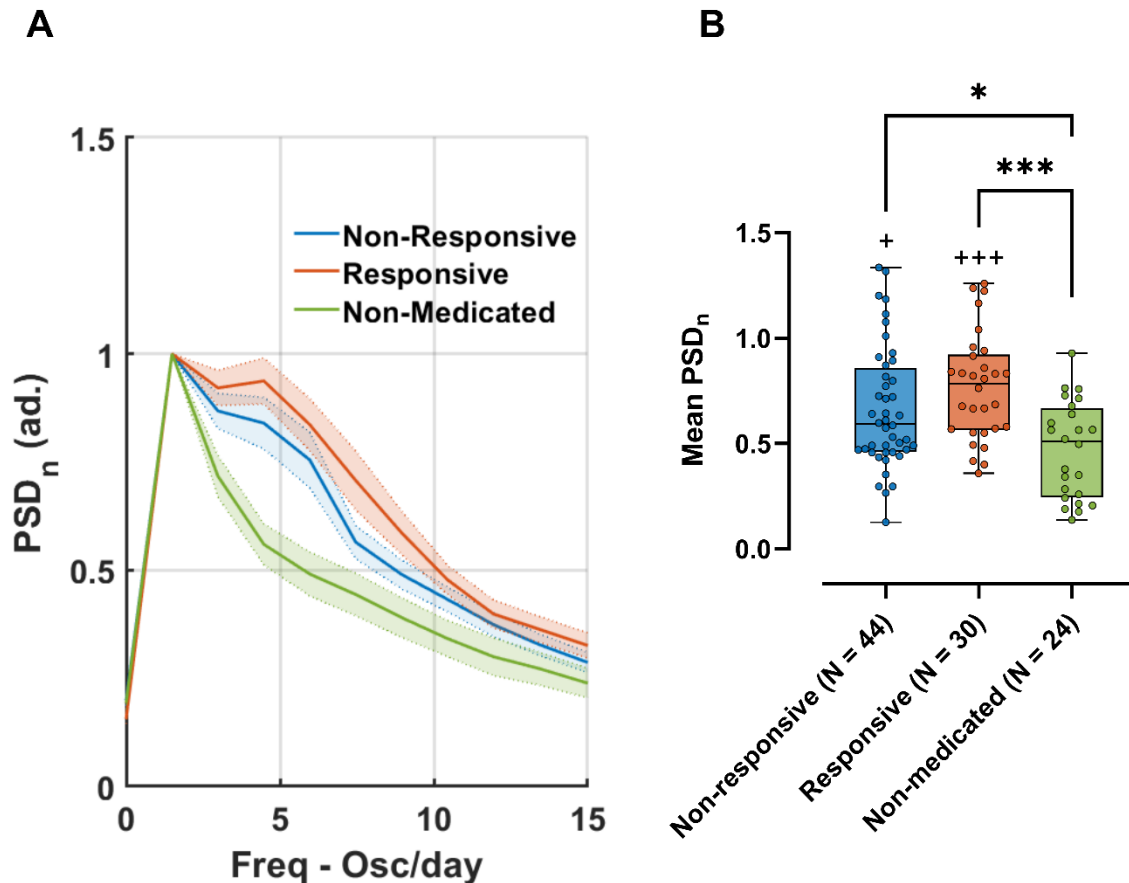
COHORT	GENDER (F/M)	AGE (YEARS)	DISEASE DURATION (MONTHS)	LEDD (MG/DAY)	H&Y	UPDRS III
NM	9/15	57 \pm 12.20	31.56 \pm 16.06	0 \pm 0	1.92 \pm 0.41	36.68 \pm 10.03
NR	12/33	64.4 \pm 7.38	39.79 \pm 18.42	486.95 \pm 246.9	2.10 \pm 0.51	39.93 \pm 12.74
R	8/25	60.79 \pm 9.26	42.55 \pm 21.1	629.69 \pm 267.18	2.06 \pm 0.35	42.10 \pm 10.37

15

16

17

18



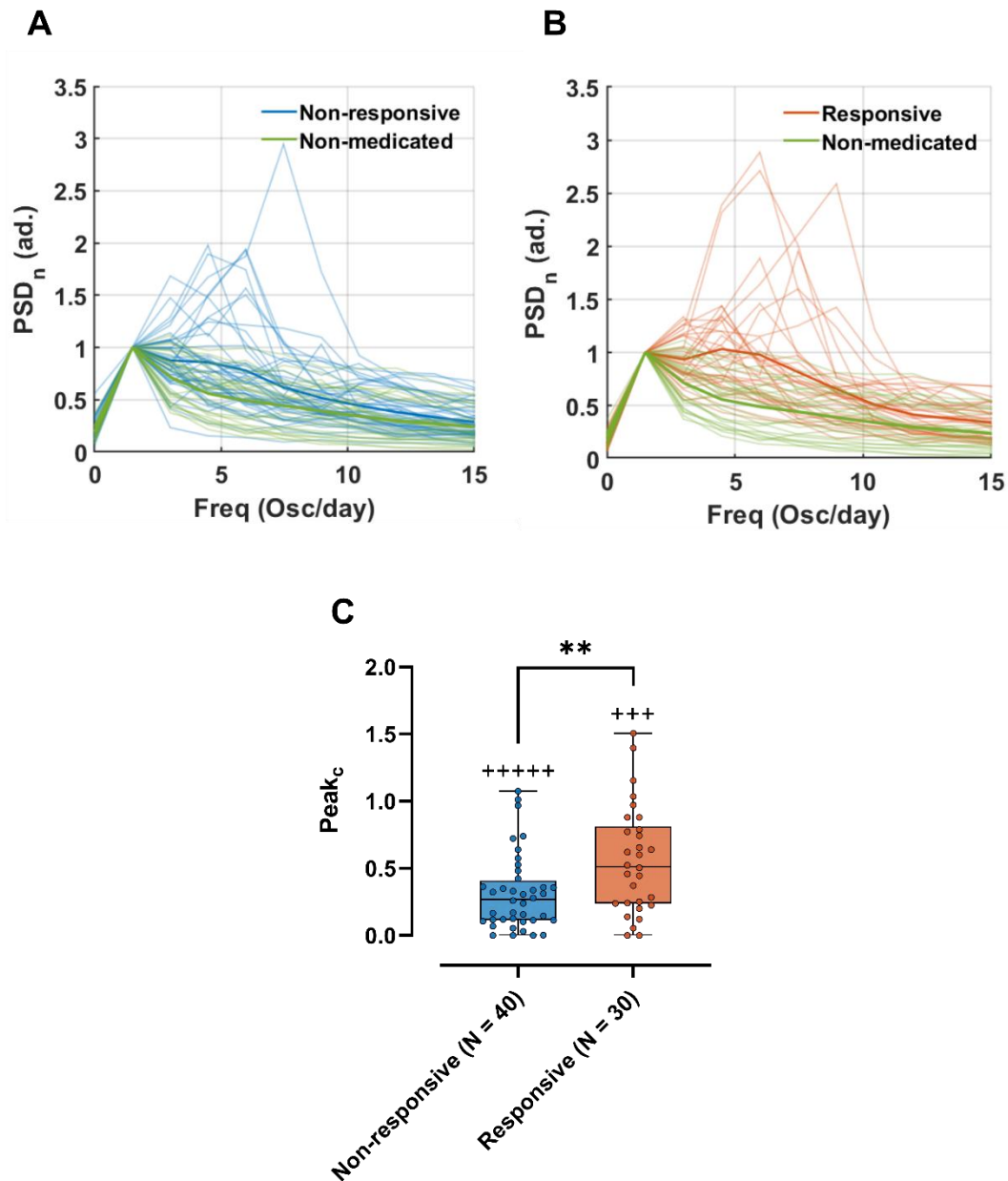
1

2 Figure S7: (A) Individual PSD_n (shaded) and group average (thick line) for the levodopa non-responsive
 3 (blue) and non-medicated (green) cohorts estimated from tremor severity time series; (B) Individual
 4 PSD_n (shaded) and group average (thick line) for the levodopa medication-responsive (red) and non-
 5 medicated (green) cohorts. Peak_c illustrates the estimation of the strength of the consistent tremor
 6 fluctuations across days based on patient PSD_n peak corrected by the non-medicated average; (C)
 7 Peak_c boxplot for the levodopa non-responsive and levodopa responsive cohorts.

8

9 Figs.S8(A) and S8(B) show the individual normalized PSDs for wrist-estimated tremor
 10 probability across the investigated cohorts, along with their respective averages (thick lines)
 11 obtained from the wrist-based severity measure. Both the levodopa non-responsive and
 12 responsive groups contained patients with enhanced frequency components clearly different
 13 from the monotonic decay observed for the non-medicated group as shown in both panels.
 14 The obtained results are like those observed in the tremor probability scenario (Figure 4 -
 15 Results). Once again, the enhanced frequency components occur between 3 and 10
 16 oscillations per day and the frequency peaks are more prominent in the levodopa-responsive
 17 cohort (Fig. S8(B)) but also present in the non-responsive group. The strength of the consistent
 18 tremor fluctuation (Peak_c) is significantly higher for the levodopa responsive group when
 19 compared to the non-responsive group (Mann-Whiney test: U = 375, p = 0.007) (Fig. S8(C)).

20



1

2 Figure S8: (A) Individual PSD_n (shaded) and group average (thick line) for the levodopa non-responsive
 3 (blue) and non-medicated (green) cohorts considering wrist-based tremor severity; (B) Individual PSD_n
 4 (shaded) and group average (thick line) for the levodopa medication-responsive (red) and non-
 5 medicated (green) cohorts. Peak_c illustrates the estimation of the strength of the consistent tremor
 6 fluctuations across days based on patient PSD_n peak corrected by the non-medicated average. f_D
 7 denotes patient's dominant frequency in the medication intake frequency range; (C) Peak_c boxplot for
 8 the levodopa non-responsive and levodopa responsive cohorts.

9

10

11

12

1 Table **S3** shows the pairwise correlations between key clinical aspects and $Peak_c$ evaluated
 2 based on tremor severity. In general, the exploratory analysis agrees with the results observed
 3 in the tremor probability scenario, in which $Peak_c$ significantly correlates with UPDRS 4.3 and
 4 LEDD variables.

5

6 Table **S3**: Spearman correlation between patients' clinical variables and the $Peak_c$ evaluated
 7 from the individual tremor severity. Significant positive and negative correlations are
 8 highlighted in red and blue, respectively. **Resp**: clinical tremor responsiveness to levodopa
 9 medication; **Wearing OFF Quest**: tremor improvement with medication - Wearing OFF
 10 questionnaire (yes: 1; no: 0); **UPDRS 4.3**: time spent in levodopa OFF state; **DD**: disease duration
 11 in months; **LEDD**: Levodopa equivalent daily doses (mg); **Age**: patient's age in years; **ArmActv**:
 12 mean arm activity probability; **PaidJob** – patient paid job activity (yes:1; no: 0). P-values were
 13 adjusted for a false discovery rate of 0.05 according to Benjamini & Hochberg, 1995
 14 procedure.

15

	Resp	Wear OFF Q	UPDRS 4.3	DD	LEDD	Doses	Age	ArmActv	PaidJob	Peak _c
Resp	-	0.22	0.25	0.15	0.41	0.12	-0.06	0.04	0.00	0.36
Wear OFF Q	0.22	-	0.26	0.02	0.29	0.18	-0.27	0.02	0.22	0.24
UPDRS 4.3	0.25	0.26	-	0.11	0.48	0.40	-0.24	0.17	0.10	0.43
DD	0.15	0.02	0.11	-	0.47	0.16	0.11	0.04	-0.15	0.33
LEDD	0.41	0.29	0.48	0.47	-	0.47	-0.11	0.15	-0.21	0.41
Doses	0.12	0.18	0.40	0.16	0.47	-	-0.18	0.27	-0.06	0.25
Age	-0.06	-0.27	-0.24	0.11	-0.11	-0.18	-	-0.12	-0.65	-0.10
ArmActv	0.04	0.02	0.17	0.04	0.15	0.27	-0.12	-	0.02	0.14
PaidJob	0.00	0.22	0.10	- 0.15	-0.21	-0.06	-0.65	0.02	-	-0.06
Peak _c	0.36	0.24	0.43	0.33	0.41	0.25	-0.10	0.14	-0.06	-

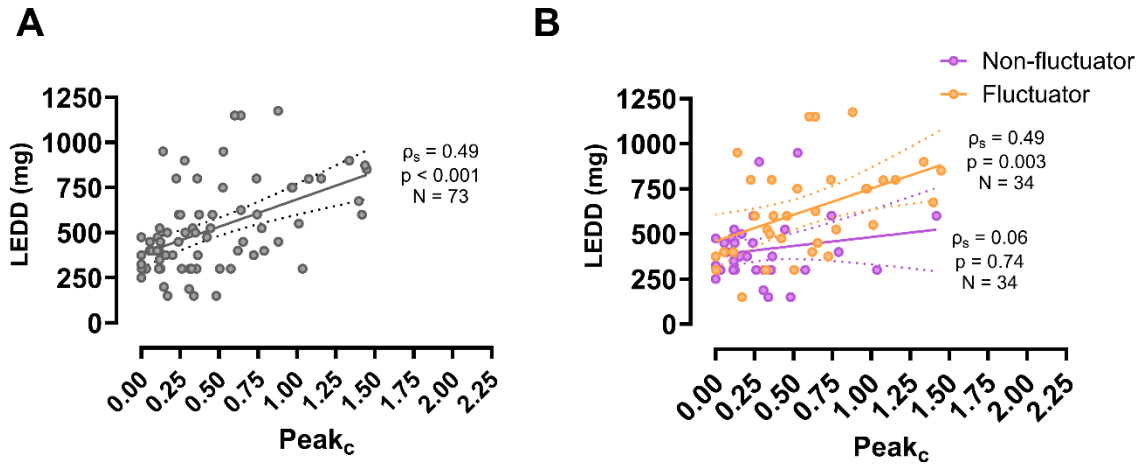
16

17 As observed in the tremor probability scenario (Figure 5 – Results), the correlation between
 18 the fluctuation strength of the fluctuation and LEDD is primarily driven by the subset of
 19 patients experiencing OFF states (symptom fluctuator cohort, UPDRS 4.3 ≥ 1) as show in Fig.
 20 **S9(B)** (cohort in orange).

21

22

23



1

2 **Figure S9:** (A) Levodopa equivalent daily dose (LEDD) vs. the strength of the consistent tremor
 3 fluctuations across days (Peak_c) for the medicated cohort evaluated from tremor severity in the
 4 levodopa constancy responsiveness scenario. Outliers were excluded for both variables. Spearman
 5 correlations (ρ_s), p-values (p) and number of samples (N) are shown in the panels. (B) LEDD vs. Peak_c
 6 dispersion considering further stratification into symptom fluctuator (UPDRS 4.3 \geq 1) and non-
 7 fluctuator (UPDRS 4.3 = 0) groups.

8

9

10

11

12

13

14

15

16

17

18

19

20

21

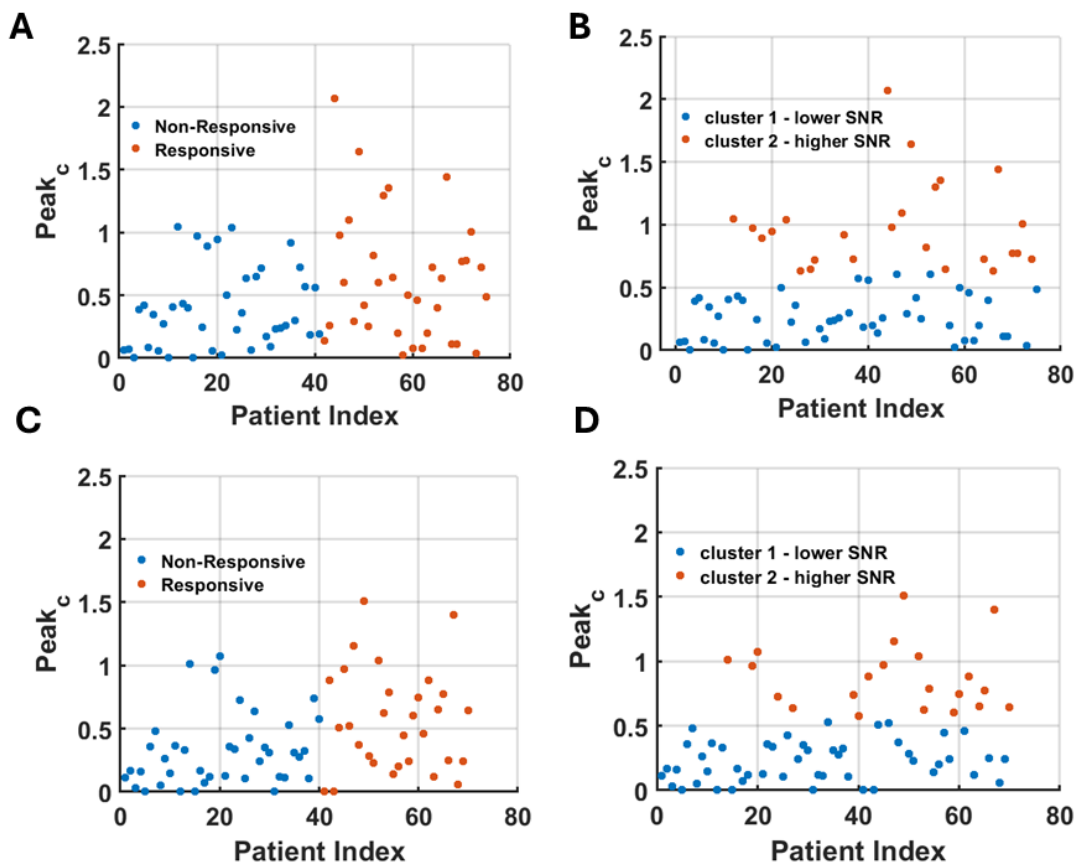
22

23

24

1 **Peak_c clustering and dominant frequency relation with the number of medication doses**

2 Figures **S10(A)** and **S10(B)** show, respectively, the dispersion of the consistent tremor
3 probability fluctuation strength (Peak_c) and its unsupervised (k-means) clustering into groups
4 of higher and lower tremor fluctuation strength (or signal to noise ratio, SNR). Outliers of each
5 group were excluded in the analysis. In this case, the dominant frequency of the fluctuation
6 activity significantly correlates with the number of levodopa doses for the group with higher
7 SNR ($\rho_s = 0.46$, $p = 0.03$, $N = 22$), suggesting a medication-dependent modulation when
8 representative tremor oscillations are present. Similar panels are show for tremor severity
9 (Fig. **S10(C)** and **S10(D)**), in which no such correlation was observed ($\rho_s = 0.11$, $p = 0.65$, $N =$
10 20).



11

12 Figure **S10: (A)** Dispersion of the patients' PSD ratio peak (Peak_c) evaluated from wrist tremor
13 probability considering the clinically rated levodopa non-responsive and responsive groups
14 (outliers excluded); **(B)** Best k-means Peak_c clustering solution (N = 100 runs) for identifying
15 patients with higher consistent tremor severity fluctuations across days. **(C-D)** Similar panels
16 for tremor severity fluctuations.

17

18